

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 7, 2004, 20:11:19 ; Search time 4424 Seconds
(without alignments)
10639.813 Million cell updates/sec
Title: US-09-993-777-1_COPY_9874_10959
Perfect score: 1086
Sequence: 1 ATGGCATCGTACTTGGAC.....CGCCGACTCAATAAAATG 1086
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1084.4	99.9	18912	14	ILU28832	U28832 Infectious
2	63	5.8	1627	14	HSMGDGLYCO	L31965 Gallid herp
3	51.2	4.7	125020	9	AF429315	AF429315 Homo sapi
C	47.8	4.4	2000	6	AX655393	AX655393 Sequence
5	47.4	4.4	163025	14	AY372243	AY372243 Peittacid
C	43.2	4.0	1246	6	AX164174	AX164174 Sequence
C	41.4	3.8	1598	8	AK119529	AK119529 Oryza sat
C	41.4	3.8	1618	8	AK065286	AK065286 Oryza sat
C	41.4	3.8	1756	8	AK121636	AK121636 Oryza sat
C	41.4	3.8	2000	6	AX655393	AX655393 Sequence
10	41.4	3.8	138025	8	AF003261	AF003261 Oryza sat
11	41.4	3.8	200183	8	AF003227	AF003227 Oryza sat
12	41.4	3.8	200183	8	AF003227	AF003227 Oryza sat
C	40.2	3.7	125020	9	AF429315	AF429315 Homo sapi
14	39.6	3.6	5513	3	AY135117	AY135117 Drosophil
15	39.6	3.6	5514	3	AY135123	AY135123 Drosophil
16	39.6	3.6	5521	3	AY135128	AY135128 Drosophil
17	39.6	3.6	5525	3	AY135131	AY135131 Drosophil
18	39.6	3.6	5526	3	AY135127	AY135127 Drosophil
19	39.6	3.6	19277	3	DMU19909	U19909 Drosophila
C	39	3.6	110000	2	LMFLCHRA8_04	Continuation (5 of
C	38.8	3.6	507	8	AK062637	AK062637 Oryza sat
22	38.8	3.6	70000	8	AF003848	AF003848 Oryza sat
23	38.2	3.5	591	8	AK107720	AK107720 Oryza sat
C	38.2	3.5	16984	1	AE006954	AE006954 Mycobacte
C	38.2	3.5	40051	1	MSGY224	AD000004 Mycobacte
C	38.2	3.5	109555	8	CNS08C90	AL732849 Oryza sat
C	38.2	3.5	324050	1	EX248335	BX248335 Mycobacte
C	38.2	3.5	342416	1	EX842573	BX842573 Mycobacte
29	38	3.5	2677	6	AX777355	AX777355 Sequence
30	38	3.5	2838	6	AX777353	AX777353 Sequence
31	38	3.5	4516	3	BT001484	BT001484 Drosophil
32	38	3.5	4543	3	DROCSW	M94730 Drosophila
33	38	3.5	5492	3	AY135129	AY135129 Drosophil
34	38	3.5	5511	3	AY135134	AY135134 Drosophil
35	38	3.5	5521	3	AY135132	AY135132 Drosophil
36	38	3.5	5521	3	AY135133	AY135133 Drosophil
37	38	3.5	5522	3	AY135118	AY135118 Drosophil
38	38	3.5	5522	3	AY135119	AY135119 Drosophil
39	38	3.5	5522	3	AY135120	AY135120 Drosophil
40	38	3.5	5522	3	AY135121	AY135121 Drosophil
41	38	3.5	5522	3	AY135122	AY135122 Drosophil
42	38	3.5	5522	3	AY135124	AY135124 Drosophil
43	38	3.5	5522	3	AY135125	AY135125 Drosophil
44	38	3.5	5522	3	AY135126	AY135126 Drosophil
45	38	3.5	5523	3	AY135130	AY135130 Drosophil

ALIGNMENTS

RESULT 1
ILU28832

LOCUS

DEFINITION

ILU28832 18912 bp DNA linear VRL 09-AUG-1996
Infectious laryngotracheitis virus US10, US2, protein kinase, UL47,
glycoprotein G, ORF5, glycoprotein D, glycoprotein I, glycoprotein
E, ORF9 genes, complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphan herpesvirinae; Infectious laryngotracheitis-like viruses.

REFERENCE

AUTHORS

Wild,M.A., Cook,S. and Cochran,M.

TITLE	A genomic map of infectious laryngotracheitis virus and the sequence and organization of genes present in the unique short and flanking regions
JOURNAL	Virus Genes 12 (2), 107-116 (1996)
MEDLINE	97033380
PUBMED	8873127
REFERENCE	2 (bases 1 to 18912)
AUTHORS	Wild, M.A.
TITLE	Direct Submission
JOURNAL	Submitted (09-JUN-1995) Martha A. Wild, Syntro Research Laboratory, 3535 General Atomics Court, San Diego, CA 92121, USA
FEATURES	Location/Qualifiers
source	1. .18912
	/organism="Gallid herpesvirus 1"
	/mol_type="genomic DNA"
	/strain="USDA challenge strain"
	/db_xref="taxon:10386"
repeat_region	<1. .2909
	/notes="large inverted repeat flanking the short unique region"
	/rpt_type="inverted"
	697..1533
CDS	/codon_start=1
	/product="US10"
	/protein_id="AAC55093.1"
	/db_xref="GI:1486485"
	/translation="MENLDGCVPLAMDSHTAHVPRGRERRQGAVASSESADRS VDCIRASRLMRELVEISELKDVGSTESARDNRNALLANERLSAFAGASRATR GLGLRPMASTSVANPTDNNNGLEGERMGEIGDFWLDSDGDFEDESRTM QSNMRFVIEKLLSLRPADLASAERETSRLLAAGHWCCLMHRPCEACLYD SIYVQLFCVGTGRVQSEWRMRRLAYLAALRAGAAAANSPEVSASIFARDGIALALAR RR"
CDS	complement(1900..2784)
	/notes="coding region unique to avian herpesviruses"
	/codon_start=1
	/product="unknown"
	/protein_id="AAC55094.1"
	/db_xref="GI:1486486"
	/translation="MSKCYCLARHLKSPRCVGRVAFGLATMSRPTSHLDLAFSA AFGTDLGGRFWRAGSCDDIFWPDLAAVVOAKAYFEGKERLGSQVADITDAHD PRIAPAKAAVAAGALGTALGCGPNGEELSKYWGKQIPRAAAWEIRDPVKPVI GPDLISFFGAARVLPVLYIRARGAHSRAHNNQSSAPAAGLAIRIGEMVRSLV IALPLNLTPELDLPGSSONSIRAFVAHLMNCVATDKIMSPDVRVPEESFYSHCLRE IINCEARFCVPCNPPKM"
misc_feature	2910..16003
	/notes="short unique region"
CDS	complement(2916..3605)
	/codon_start=1
	/product="US2"
	/protein_id="AAC55095.1"
	/db_xref="GI:1486487"
	/translation="MAPVKVTIIVSAVDSHYKLPNSRFLSDSGWKELVHAVKTMAS YD RPTSLVIVRPAASLYVSGELFSLPRMCPVIRFGSGDPPGSPWSEWGLDAGFYHLS SGAYAKEFHLWLTADICMAALNPAPKTLITETGKNTERGVEIFLVNGDKTTL SLGHPSWTTLTASSLRTWPFTVTFVKVFNKSAAYCYSDSDGGERQPKFGLSLFK SKKPSRRRR"
CDS	3694..5124
	/codon_start=1
	/product="protein kinase"
	/protein_id="AAC55096.1"
	/db_xref="GI:1486488"
	/translation="MRFRICSRSAEKRRRTTENPLTSKRVCVLDLSFRTWSLRPYA EILPTAGVERLAELSVTTERAEFVTENTAVNSIPANENGONFAYAGDGFSTTEK VDGSHTDPRDASDAGPVLATRLKHSDFLQHFVRLDLDVEGAYGICDVRVYTE BEGRRGVNSNGSKCKELIAKYKNGTRASOLENILLVRLNHNVLKIOELL RYQNTTMYLQTFDYSYNDIADFDKDSFMLKTRIMKQLMSAVYIHSKLIH RDLKLENFILNCKGTVLGDFGTVPENEREPEFYGVGTVATNPFELKADSYCEI TDWCSQVWLEWYVSEFCFIDGGGNPHQQLKLVDSLSCDDEFPDPPONLYLH YASTDRAGHTVPSLRNLHLPADVEYPLVKQLTFDWRLRPSAEVLAAMPLFAEBERT ITIIGHGKPIREIRARVPSNSEG"
CDS	5210..7081
	/codon_start=1
	/product="UL47"
	/protein_id="AAC55097.1"
	/db_xref="GI:1486489"
	/translation="MTLPHRLTKPRPARFCSPVFVIHSETKLDRYNKTMLLXRPDST MRHSGGDANHRGIRPRKSGIFASAREKTKGNALTESSSSMDLDPFSTKDFGKGW TVDGPADITAEVLQAWDLQVVKHEDAEERVTYESKPTIQPFNANPDGFPWNAQD FTRAPIVYPSAEVLDAELVGAFAVSRLVQCVFTSRKSKSVTRQAQSLGDSFWRIM QNVYTVLROHITRLRHPSSKIVNCNDPLWYANQFHWGRVPSLKLASPEEYNI QHGPMAVFRNAGAGLFLPAPMAAEEERDKLRLACLASLDMDAVLAASFPYWRG VQDTSRFPALGCLSYFALVLLAETVLATMFEDHALVFNALADGDFDDYDTRID PVNENLNGAEGTLRGIIVASNTALVAVCANTYTIKLPVSATSNVAIRVETTLKA RRPQMSDIYRILQKBEFFVIAMQVRVATHANFCINTLKRSVDTGAPPLFRASSEKRL QOLNMLCPLLVIQYEDSFKAMGSELKREKLETFVKATSSDRDPGRSLFLSDHAR EIIADGVREKPVIDEFVRASVALSTAAAGKVKARLRTSVRAPVPGAGAVSARKSEI"
	7245..8123
	/codon_start=1
	/product="glycoprotein G"
	/protein_id="AAC55098.1"
	/db_xref="GI:1486490"
	/translation="MSGFNSIGSIATVSLVCSLLCASVLGAPVLGLESSPPFPFGGKI IAQCNKRTTIEVTPWSDYSGRTEGVSVKVFYGNPNPESFVGVDSETSGSHEDLS TCWALIHNLNASCVRSDAGIPDFDKCEKVORRLSGVELGYSVNGSLVLYPGMY DAGIYAYQLSVGGKGYTGSVYLDVGNPGCHQDQGYTYTSLADEASDLSSYDVASPEL DGPMEEDYSNCLDMPPLRPWTTCVSHDVEQENATDELVLWDEECAGPLDEYVDERSE TMRMVVPSPESTLQ"
	8333..11290
	/note="ORF5"
	/codon_start=1
	/protein_id="AAC55099.1"
	/db_xref="GI:1486491"
	/translation="MGTMVLVRLFLAVADAALPTGFCRCVWVVPVGGTIIQENLAVLA ESPVTGATVPPPEGAVSQIFATDPLIRYCATDEBELALERTSADADNTFSL YRPRPEIHGAYFTIGVFATQOSTESSYVSRVLNVALSRLSVRTPCCDNLQNEF TWGSKRWLGPSPVVRNDVAVLTAKAYIGECYSNSAAQTGLTSLNMTFFYSPKRYN VYTTGQSPSRITVYSSRENGOPVLNRVSDGFLVKYTPDIDGRAMINLVANTYSPADS GSULVATAPREGKLPSAIOLRHIDMGTEPPGTETTPDCOKMLETPYRGLSGNVPRD SIRGALPFPDPAADFDGTSTPTTPEPAITLIPRSTDMGFESTARAGSET LSVPEQTDRLTTLPLTPGSESENTLFPPTAPGISTGISTEPAAHETTQSAETV VEQSPSESETARSQSEBPYFTQSTEQAAALTOIAETALFOTPSAEQMTFT QPKWTEAAQTPSTIPELTQSTPPTAPSAAPVFTQSSSTVTEVFTQPS TVPMTLSSSTPAITFRTOSAGTEAFTQSSAEPDMTQSTETHEFTQAPSTVPA QTPSTBPEVLITQSPSTPEPFTILGAEBEITQTPSAPEVYTRSSSTMPETAOST LASQNTSSGQTHNTEPRTVPTTHTOKLTKENKLEFPFTVVSSEHFMSTAESOT PLDVKILVYKFSNDGTEATCVTSKVSRYVEINMVDLVNDVEISGNSPAGVFN S KNAQKQILYRVTDGRTSVQMLCSLSTSHSPPEPYCLDFTSLIAEKKDIAPELYTSDP QYATCTLTPSGVVRPEWLNNSVLPYLTTATVTVSHTAGQSTVMSKSSAGAEANIS GRGNIYECTVLISDGRVTRKERCITNTWIAVENGAAQAQLYSLFSLGSLGCGSI SALYATLWTATYF"
	11098..12402
	/codon_start=1
	/product="glycoprotein D"
	/protein_id="AAC55100.1"
	/db_xref="GI:1486492"
	/translation="MHRPHLRHSRYAKGEVLNKMDCGKRCSSGAAYFTLPWTCV RIMEHICFVRNAMDRLFLRNAPFTIIVLSSPASQSTAAVYDIILGRALDALITP AVGNRYLTVRSRCDVNLNPIINVDMDISAAKEKKGPPFEASVYVYVYIKGDDG EDKCYPIVREYREGDVLSECAQAWADVVPSTLVSRNAGAGLTFSPSTAAL SQYLLTLKGRPAQTALVTLVNDRLCKIGSOLNLELPKSCWTTTEQYTGQGHLYP IADUTNRHADVVINGYEDILQNNLNRKNPSAPDRPDSVPQEIIPAVTKKAGRTP DAESEKKAPEPESDDMAEAGENFAALPEDEVEDPEHDDPDPPYNDMPAV IPVEETKSSNAVSMPIFAAFVACAVALVGLLVMSIVKCAR"
	12510..13598
	/codon_start=1
	/product="glycoprotein I"
	/protein_id="AAC55101.1"
	/db_xref="GI:1486493"
	/translation="WASLLGTALLAALAPFGAMGIVITGNHVSARIDDDHIVIVAP RPEATIQLOLFFMPGQPHKPYSGTVAVFRSDITNOCYQLBSERENCTHRSSVFP VCKVYTSASNLGPPPEPKLITRNPRNDSGMEFYIVRLDDTKPEIDVDAIOL SVYQFANTAATRGYSKASCTFGCLPTQVLEAYLTRSNRNQAYATATATTSAA TTPPTVATTSASELEAEHFTFPLENGVDHYEPTPANESNVTVLRTMGFTLLGVTV AAVVSATIGLIVISIVTRNMCTPHRKLDTVSQDDEBSQTRRESRFGPNVACEINK GADODSELVELVAIVNPSALSPPDSIKM"
	13792..15291

```
/codon_start=1
/product="glycoprotein E"
/protein_id="AAC55102.1"
/db_xref="GI:1486494"
/translation="NMMLVILASCLARLTATRHVLFLEGTQAVLGBDDPRNVPBGT
VTKMKVLNACMKAAADVCSSPNVCFHDLIVDGGKDCPPAGPLSANLIVLLKRGES
FVVLGSLHNSNITNMTEYGGLLPDPVTRSDGIIYPRISQPDLANETTSYNVSVL
SHVDKAPHEVEIDITKPSAHARHVELQMLPFHLDNPSFTYVPLRVFPPTFHV
KNFTYSYFPFVDECEBEVKLFEPVYHPTDCKQFPATNQCILIGSLVMAEPFLGAA
SLLDGSRDLEBCHENRPNLFDRLSRLSAGLVISPLIAIPKLVIIIVSDGILGW
SVTLGKRNPRVVAETMPSKVPNMKVIGSPGMDTGNVXMYFVAGVAATCVIL
TCALLVKKKCAHQMGFSKTEPLIYAPLPKNEPEAGGILTDDEEVIYDEVYELFRGY
CKQEFREDVNTFFGAVVEGERALNFASALASWADRILANKSGRRNDSY"
CDS
15298..16080
/note="ORF9"
/codon_start=1
/protein_id="AAC55103.1"
/db_xref="GI:1486495"
/translation="MFFKTRGAEDAAAGKPRFKSRNREIILPRLGRTGKKTAGLSNY
TOPIPWNPFCARGESDNHAKDTFYRRTCASSTVSSQPDSPHTPMPTPEYGRVPS
AKRKLSDSCDGHQPIVSKCLPDSQAPARTYSYSAORYTDEVSSPTFPGVDVAD
LETRALPGATTQETESKNKLPNQSRLKPKPTNEHVGERCPBSGTVEAFSLGILSR
VGAATANELARMERACPLAASAAAAGIVANAAARALQKGR"
repeat_region
16004..16912
/note="large inverted repeat flanking the short unique
region"
/rpt_type=inverted
16125..17013
/note="coding region unique to avian herpesviruses"
CDS
/codon_start=1
/product="unknown"
/protein_id="AAC55104.1"
/db_xref="GI:1486496"
/translation="MSKCYCLARHLKSPRCVGRVAVGGGLATMSRPTSHLDLAFSA
AFRGDLPGRFWRASQSDIFPFDLAAVIVQAAARVFEKERIGSLQVARDTAHD
PRIPAARAAVAAVGLTALSELVPGNGELESKWGKQIPRAAAHIVRDPKVPVI
GPDILFSFAVELPVLVYIRARGGAHSAHWNQSSAPAAGLAIRIGMEWRSLLV
IALPLSNFLPDELPGSSONSIRAFVAHLMNCVATDKIMSPDVRVPEVESFYSHLRE
IIMCRARCFPCNPBPWK"
complement(17380..18216)
/codon_start=1
Query Match 99.9%; Score 1084.4; DB 14; Length 18912;
Best Local Similarity 99.9%; Pred. No. 2.1e-290;
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCATCGTACTTGGAACTCTGGCTCTCCTTGGCGGACGCTCGACCCCTTCGGCGCG 60
DB 12510 ATGGCATCGTACTTGGAACTCTGGCTCTCCTTGGCGGACGCTCGACCCCTTCGGCGCG 12569
QY 61 ATGGGAATCGTGATCACTGGAATACACGCTCTCGCGCAGGATTGACGAGATCACATCGTG 120
DB 12570 ATGGGAATCGTGATCACTGGAATACACGCTCTCGCGCAGGATTGACGAGATCACATCGTG 12629
QY 121 ATCGTCGCGCTCGCGCCGAACTACAATTCAACTGCAGCTATTTTTCATGCTGCGCCAG 180
DB 12630 ATCGTCGCGCTCGCGCCGAACTACAATTCAACTGCAGCTATTTTTCATGCTGCGCCAG 12689
QY 181 AGACCCCAAAACCCCTACTCAGAACCGCTCGCGCTCGCTTGGCTCTGATATAACAAC 240
DB 12690 AGACCCCAAAACCCCTACTCAGAACCGCTCGCGCTCGCTTGGCTCTGATATAACAAC 12749
QY 241 CAGTGTACAGGAATCTAGCAGGAGCGCTTTGAAAATTCACATCATCGATCGTCTTCT 300
DB 12750 CAGTGTACAGGAATCTAGCAGGAGCGCTTTGAAAATTCACATCATCGATCGTCTTCT 12809
QY 301 GTTTTGTGCGCTGTAAGTACCGAGTACAGCTTCTCGCTCGCTCGAACHAGCTAACCGGA 360
DB 12810 GTTTTGTGCGCTGTAAGTACCGAGTACAGCTTCTCGCTCGCTCGAACHAGCTAACCGGA 12869
QY 361 CTTCCACACCGCTTTAAGTCTACTATACGAATCTCTCGTCCGACACGACGCGGATGTC 420
DB 12870 CTTCCACACCGCTTTAAGTCTACTATACGAATCTCTCGTCCGACACGACGCGGATGTC 12929
```

```
QY 421 TACGTAATTTGTTGGCTAGACGACCAAAAGAACCCATTGACGTTCTCGGATCCAACTA 480
DB 12930 TACGTAATTTGTTGGCTAGACGACCAAAAGAACCCATTGACGTTCTCGGATCCAACTA 12989
QY 481 TCGGTGATCAATTCGCGAACACCGCGCGACTCGCGGACTCTATTCCAAAGGCTTCGTGT 540
DB 12990 TCGGTGATCAATTCGCGAACACCGCGCGACTCGCGGACTCTATTCCAAAGGCTTCGTGT 13049
QY 541 CGCACCTTCGGAATACCTACCGTCCCACTTGGAGCCCTATCTCAGGACCGGAGGAAGTTGG 600
DB 13050 CGCACCTTCGGAATACCTACCGTCCCACTTGGAGCCCTATCTCAGGACCGGAGGAAGTTGG 13109
QY 601 CGCAACTGGCAAGCTACGTTGCCACGAGGCGCACGACGACCGCGCGGCGGACAAACC 660
DB 13110 CGCAACTGGCAAGCTACGTTGCCACGAGGCGCACGACGACCGCGCGGCGGACAAACC 13169
QY 661 CGGACGCGCTCACTGCAACACGAGGCTCCGAACCTTGAAGGGGACACTTTTACCTTTCCC 720
DB 13170 CGGACGCGCTCACTGCAACACGAGGCTCCGAACCTTGAAGGGGACACTTTTACCTTTCCC 13229
QY 721 TGGCTAGAAAATGCGGTGGATCATTACGAACCGACACCCCGCAACGAAAAATTCAAAACGTT 780
DB 13230 TGGCTAGAAAATGCGGTGGATCATTACGAACCGACACCCCGCAACGAAAAATTCAAAACGTT 13289
QY 781 ACTGTCGCTCTCGGACCAATGAGCCCTACGCTAAATTGGGGTAACCGTGGCTGCGCTCGTG 840
DB 13290 ACTGTCGCTCTCGGACCAATGAGCCCTACGCTAAATTGGGGTAACCGTGGCTGCGCTCGTG 13349
QY 841 AGCGCAACGATCGGCTCGTCAATTGTAATTTCCATCGTCAACGAAACATGTGACACCCCG 900
DB 13350 AGCGCAACGATCGGCTCGTCAATTGTAATTTCCATCGTCAACGAAACATGTGACACCCCG 13409
QY 901 CACCGAAAATTAGACACGCTCTCGACGACGAGAGAGAAAGTTTCCCAAACTAGAGGGAA 960
DB 13410 CACCGAAAATTAGACACGCTCTCGACGACGAGAGAGAGTTTCCCAAACTAGAGGGAA 13469
QY 961 TCGGAAAAATTTGACCCATCGTTGCGTGGGAAATAAACAGGCGCTGACCAAGATAGT 1020
DB 13470 TCGGAAAAATTTGACCCATCGTTGCGTGGGAAATAAACAGGCGCTGACCAAGATAGT 13529
QY 1021 GAATCTTGGAACTGGTTGCGATTGTTAACCCGCTCTCGCTTAAGCTCGCCGACTCAATA 1080
DB 13530 GAATCTTGGAACTGGTTGCGATTGTTAACCCGCTCTCGCTTAAGCTCGCCGACTCAATA 13589
QY 1081 AAAATG 1086
DB 13590 AAAATG 13595
RESULT 2
HSMGDGLYCO
LOCUS
DEFINITION Gallid herpesvirus 1 glycoprotein D (gp) gene, complete cds.
ACCESSION L31965
VERSION L31965.1 GI:493595
KEYWORDS glycoprotein D,
SOURCE Gallid herpesvirus 1
ORGANISM Gallid herpesvirus 1
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
REFERENCE 1 (bases 1 to 1627)
AUTHORS Johnson,M.A., Tyack,S.G., Prideaux,C.T., Kongswan,K. and Sheppard,M.
TITLE Sequence characteristics of a gene in infectious laryngotracheitis virus homologous to glycoprotein D of herpes simplex virus
JOURNAL DNA Seq. 5 (3), 191-194 (1995)
MEDLINE 95337426
PUBMED 7612933
FEATURES
source
1..1627
/organism="Gallid herpesvirus 1"
/mol_type="genomic DNA"
```


RESULT 4
AX655393/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Query Match
Best Local Similarity
Matches
40; Conservative
254; Mismatches
241; Indels
0; Gaps
0;
390 AAATCTCGTCCGACGAGCAGCGGATGTTCTACGTAATGTTGGCTAGACGACACCA 449
Db
854 WMMWRYTYCYAMTCAKCKYKAMTKMTTWACAWRATSRWRAMAGMRKRYEMKRA 795
Qy
450 AGAACCCATGAGCTTTCGGATCCAACTATCGGTGATCAATTCGGGAAACACCGCCG 509
Db
794 YMMWRWRCWAGWMMKSRKRWKWKYATRYKWKWMTWMMWRWMSYRWMSGWM 735
Qy
510 GACTCGGGACTCTATCCAGCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 569
Db
734 RWSAWRYCRRMKCAKTKYASSARFWTKRAKRSYRFRWYKRGWTKYRYRWSRMTTRA 675
Qy
570 TGAGGCGTATCTCAGGACGAGGAGAAAGTTGGCGCACTGGCAAGCGTACGTCGACGGA 629
Db
674 RMRKRKAGASWKSWMYWRGARSWMYSKYSAKCKCKTRMTSSYMSGTGMYMSY 615
Qy
630 GGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 689
Db
614 KMSWTSKMSYMGKMTCTMTYSKSTRSKWSGMSMYRWKWKYKRYMYK 555
Qy
690 CGAACTTGAGCGGAGACATTTTACCTTCCTCGCTAGAAAATGGCGTGATCATACGA 749
Db
554 KCTWRRWMCYRWGTYMTTTSRSMYTYGRYKARYTSKRMYMYKRYKCYWYTYGMYMK 495
Qy
750 ACCGACACCGGAGAAACGAAATTCAGCTTACTGTCGTCGTCGTCGTCGTCGTCGTCGTC 809
Db
494 SYMMRYGYCKACCKCCYACWKAAYSGMMYMYTKYSKWRWMSMTKYMWSMYKCKRSKY 435
Qy
810 GCTAATTTGGGTAACCGTGGCTCGCTGCTGAGCGCAACGATCGGCGCTGCTGCTGCTGAT 869
Db
434 GAGCTGCKRWYTCSTGIMKWTYMGYSKYRSCTKYNRMIMYKWMYMYTAYSSMTW 375
Qy
870 TTCATCTGTCACGAGAAACATGTGACCCCGCAGCAACGAAATAGACACGCTCG 924
Db
374 YYYAKYKWKYKRRGTMSYKYSKYKCTWYKWKYKWKYKWKYKWKYKWKYKWKYK 320
RESULT 5
AY372243
LOCUS
DEFINITION
ACCESSION
VERSION
AY372243
Paittacid herpesvirus 1 isolate 97-0001, complete genome.
VRL 13-SEP-2003

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS
gene
CDS
gene
CDS
product="DNA helicase primase component"
protein_id="AAQ73688.1"
db_xref="GI:34500205"
translation="MFGLIGVFDRREESKCEPLVKHAPAIPEASEAENYPLSSA
DDKTLAALDAVETTLIGSCSLHVTDTAAWPPAGDSAAAAMVMMGSHSPGRLMIEAL
ESVRSALGTHSDATGSCQRCISALAPQSPQSVPEGPINDSGQIFAVSAGAGEENWI
DDEALCAADLQIDPIGDAQSVGGSTAACLPSPQKEERILITATDGYFVCSHLILL
GQALLQELVYVAYSHVTPSAGGGLRATFYLLAATGRCQRQRPVPCVFTGGA
QKTRALRGKPLWQDVIAISDAQSLAHEDLITAGCTICAGTAAAPGSDERN
LSYNYVASTPLGLRTIASLWQYEVVAAFRKYNASFCPFWFYKFGPSETSLV
AVLYLLGATQDDGIAGSDLGQKDLATFPKPGNASGLTAARLNTAFERSFWC
CSELAGPKABKIFREYDKTRVSSDTLLYQSLDSLIEKYRHALLPAXDFVRFVYLAVH

EGNRAAEVMEALRNGGRARTATTAQSBFFKGLRECMSPYFARNVDVAVKRL
PVRVAEHLRSYGEIPRSGFWGFCWPSYDVSRLINKI CAGPLCCLGTLQADAVRA
PPDGKGTCAETAATIIHDSASGESEATVKCGAHMFRANKRAGDGTIRISPAKNE
ADTRAPLARTALAYATQPRGIMPVTLLLSLDTSAGDADVSEAKAQRPPVTV
RVAMPGGGAFAVADQDKTVTVDFVSGFWAPKSPESNGPLSGDADAADDAWT
RALLARWCTSVARCAANVASQVLYLNRNEVLSLAVCNPLDLDITLKSSVGRP
SMULHEMVSREALVLLWLSLFADEIDPTI PVIFYKTQCDAAAGDEEASPORRG
VGGGDDVAMGEEDIEEMCGGYYSQTAMNDMTWDDAPQMDIDSDHPAEPMDD
GAGAAIEBKNTAAAFACRICKIFGRICVPIPKPYALAGLSVAKATSLAQVQATVLQ
EKFEVACHADPADIYVDFGVSGVAGSLRPLFFAKVNRGVNMLPLVFPKPCRLA
RSEFVACHADPNFHYGAARONPAPLTIITSTVTRDVLADONRAHGKPRAGTARLA
DALAAVGVPGNQOCTPGDACAVDASWLLDSVAMPALREYIDIEHPFRAAEVRDARTD
CIVYGRICACALRGGAGYNGRRSNFTCLKYQHRGASNQTVIASVVVAVNSQGLPY
AAIQTCFATKCSNELQFTVTLSATKDV"
/gene="UL101
/gene="UL51"
10319..11101
/gene="UL51"
/codon_start=1
/product="putative"
/protein_id="AAQ73689.1"
/db_xref="GI:34500206"
/translation="MRNP1FLALYTGSHWIRQMPMAQCLKSCVCSYBPLVDVA
QLPVDDEGRVTAANVVKLPDCLSDVVLKSDDELKLAQARNISKILAKSATAIH
IARGIKCPKTEVLRQTLVDNCFKSVSYSLAYLYLSPGADTDMIDQVVAQTADR
TIMLGMVTLSHAMRVGDVDI PHSODLAKMGLSAADSDTLDPIQLVGWMEPPSPPH
TADARDAHVSPERRPOSSATLTSIDAKSRGLAQVPAS"
complement(11232..12476)
/gene="UL50"
complement(11232..12476)
/gene="UL50"
/codon_start=1
/product="dUTPase"
/protein_id="AAQ73690.1"
/db_xref="GI:34500207"
/translation="MEATTKOAMIEIGKWAASCLARTCVISNEEDIYABRADTP
VLKDSVTALPGVGIHVSIGTARAKTAWELVGLVDSQVGTGLLGLTDTPTA
TSGAGGIVSFGGVHARTVTKLVDPDIIMGACGAGARLPKTAITLTFKGDDELLG
NGRDHOMESLASIYPLILVQLDPCVYFGCTGCKAFYRLGSLCTEPLNELGSHIY
LRGSAYESVFAADPDMFVAMYGKLLIGMAETPNEKLTVELRDDSDSPAALIFP
HDTFGKEADEAGDYIRAPENCTLPFGSVVRLKQKLMHGKRAAFVNGRSMNLK
VLVEPVRVDDEWVSFNTIIRDAAPFRKNDRIAQLVALEDKLELMGVDALPWRVV
QSVQEEKNSRGDKFGSGGV"
12667..13014
/gene="UL49.5"
12667..13014
/gene="UL49.5"
/codon_start=1
/product="glycoprotein N"
/protein_id="AAQ73691.1"
/db_xref="GI:34500208"
/translation="MWLLRPAGSNFTVALIVLACAGPLTCSAQLDAGILNPNWSAGHN
DAYPMGFANSESDERFYPHPCSSRGLPLVNESMASVIFFLSLAMVCVAIVAILYNCC
FNSFKNSVINSRW"
13184..14038
/gene="UL49"
13184..14038
/gene="UL49"
/codon_start=1
/product="vp22"
/protein_id="AAQ73692.1"
/db_xref="GI:34500209"
/translation="MSYVTEQSGEKKTRKSSSRKSDRSASSSPPPPGAVRGVRV
MYPPGVDGAWLRRRRRSGSSSDSSRRDDRRQPEQOPKQOSTRERKRSOTTV
TTRKTKDRGGKSGNSGNSPISDNCANLVKLSLAKGFSATPASPASDRWRITTN
PANSRFAVQSVTVMAQBELAARAWEKYKERNEDLEKLVETLEIKLTVNPGGLW
DVASNTASAIRNQPIITHDLLYASSPAGDARTSRQSKSRKSMRPRGDEDERDAGSP
RPPSSRRR"
14500..16323
/gene="UL46"
14500..16323
/gene="UL46"
/codon_start=1

/product="vp11-12"
/protein_id="AAQ73693.1"
/db_xref="GI:34500210"
/translation="MSTKSSPARGTVVIPHAMTLRSVELKKIPIKVLTVFENAQKI
SPWLSRQASLNSLAATAETMLPGIILGPSRAEHFNGDPTIPDKMVRVMEVGHKML
FRATLDDTYHLNTRIIEVREKFORARAMYTRASMLMDVELPFKDKMGEVVAADL
EATCMHWFPHTVPAPNKSDDKAYLTSVAQFARSTAFIESYMYMHPHDEPVNS
TDTNRVAAMLIIYHNGCKLASLFDALDAILRSKFNAGSRHLSGIIYGFTRNGKA
TMAAKSGNVVEQIRETCFPGSDAMVECLISALDRMTYLVARWELMSVYKFGRTGP
KAGLAWLNLNGLHSYLTIMNTVYVYLPHYLVNPLIITSSRVLGLQLCTSHT
TFENATCPMLLAQNAVEWKQBERGTAAMDVVCARAMLRASRENKRPANITAMNNT
LEEISAVLXILBEAPQAPASAPSSLCATNYCPDSPSPRPVPOKRAKDPGRFSR
IMAGDDRRASVQSRGSGYWCHEASTASTKLRGRVLSAPLIHRRKSAEGLQGA
LRGSLRLPNGGRAALPSPSPSSSESTSPPEARPETDHK"
/gene="UL45"
complement(16521..17426)
/gene="UL45"
complement(16521..17426)
/codon_start=1
/product="membrane protein"
/protein_id="AAQ73694.1"
/db_xref="GI:34500211"
/translation="MRNNGARTQTPPPAAEGTADVHNGEATIIOLDGQPSAPSP
LTVFEATDVKARRNLESEPEFVRALIDDELSDIPVAAPLDAVQKISCCGCVCL
RRCFAVLTGMAGVMLAALATIAFAVPSVMTGVGVCERGWQHVGVCKPGRTPP
PSSAADNGFAATPAGANTCAAAGSIVSDQALSFLLSKTASGNSAPRPGSW
TTHSGCAIYHVAPSSLLEVYNTSSLSQSLALRTRDALDVSIADTQCSGAGVWCAA
APAPTALKYAQLRALITLGVRAE"
21070..22260
/note="ORFA"
/codon_start=1
/evidence=not experimental
/product="hypothetical protein"
/protein_id="AAQ73695.1"
/db_xref="GI:34500212"
/translation="MDLAGLSALAKYAVLASEGLTQLVWTERDRISMGVPLNPECG
LIPCTPFAVYDRIIRCPDRDLDMRAHELTIGRTYWSLYSEFLENTWECIALP
FVPCRLGWKEAPPLASMERVASEVELGDRVTCNSVRFAVIMGMDVITPEGVSI
RWSGTGPFVFESEDACAEVVERCRESFDISCVVPLSTAPMTGNLFLTCILCVG
PRMTNPAVRGAEVLAKAIMQHTNSGRSFCNCRMETRDTPIGPRVPGSGSIFLS
PEEKAEIARNGRLSLTDKVELVDGSRKPAVISQSWARKRVRGVLPASVRLQ
TDLADGAAAGRFLQCYCYTVTRQLGFNERDARLRLRGGPVYVTTAPRGVVFDETT
DPLR"
22726..23829
/note="ORPB"
/codon_start=1
/evidence=not experimental
/product="hypothetical protein"
/protein_id="AAQ73696.1"
/db_xref="GI:34500213"
Query Match 4.4%; Score 47.4; DB 14; Length 163025;
Best Local Similarity 50.2%; Pred. No. 0.16;
Matches 145; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 156 CGAGTATTTTCATGCTGCGCAGAGACCCCAACACCTTACTCAGGAACCGTCCCGGT 215
Db 144244 CGCGCTCGTCTTTCTCGAGACCGACCCCAAGACCCATACCGGGGCGACGTAAGGT 144303
QY 216 CGCGTTTCGGTCTCATATAACAAACACAGTGTCTACAGAACTTAGCGAGGAGCGCTTGA 275
Db 144304 GCTTTTCCAAACCGCGAGTCCAGCAGTGTCTCATACCGTCTTCCAGGTACGTACTC 144363
QY 276 AAATGCACTCATCGATCGTCTTCTGTTTTTGTTCGGCTGTAAAGTGACCGAGTACAGTT 335
Db 144364 TAACTGCACAAACAGCTCGCTCGGCTTCTTCGGCTGCTTACCGCACCGACCGAGTT 144423
QY 336 CTCGCGCTCGAACAGACTAACCGGACCTCCACACCGGTTTAAAGTCTACTATACGAATCC 395
Db 144424 CTCGCTGCGCGCGCAAAACAGGGGTACGCTCTCCCGGCTTCG---TATCACTGAGGACCC 144480
QY 396 TCGTCCGAAACGACGAGCGGATGTTCTACGTAATGTTCCGGTACACGAC 444

Db 144481 CACCATGCTGCAGCGGCACATCATCTGCTGACCTCGATC 144529

RESULT 6

AX164174/c

LOCUS

AX164174.1

DEFINITION

Sequence 4 from Patent WO0138564.

ACCESSION

AX164174

VERSION

AX164174.1

GI:14545112

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1

AUTHORS

Rouleau, G.A., Lafreniere, R.G., Rochefort, D., Cosssette, P. and Ragdale, D.

TITLE

Locs for idiopathic generalized epilepsy, mutations thereof and method using same to assess, diagnose, prognose or treat epilepsy

JOURNAL

Patent: WO 0138564-A 4 31-MAY-2001;

MCgill University (CA)

FEATURES

source

1. .1246

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 4.0%; Score 43.2; DB 6; Length 1246;

Best Local Similarity 12.7%; Pred. NO. 1.2;

Matches 39; Conservative 141; Mismatches 127; Indels 0; Gaps 0;

QY 65 GAATGCTGATCACTGGAATACGCTCGCCGAGATTGACAGCATCATCTGTCG 124

DB 445 SHEKSDSSSYMYKAAACNHNATMBCCSABCSBNCBHGHBASDKYCNBTYGSY 386

QY 125 TCGCGCTCGCCGAGAGTACAACTCACTGAGTATTTTCATGCTGCGCCAGAC 184

DB 385 SASNNGNYCYDYBESHYNSHSDHHTNSCHMTYCYSAVSNNYSSCYSDSSRYMAY 326

QY 185 CCACAAACCTTACTCAGAACCGCTCGGCTCGGCTCGGCTCGATATACCAACCACT 244

DB 325 NCSYCMYSHSHHMCSSMYMYNYNTSSNSMTSSSHSYCTTSYSTATATTTT 266

QY 245 GCTACCAAGCACTAGCAGGAGCGCTTTGAAATTCGACTCATCGATCGTCTGTTT 304

DB 265 MKTWTATNRKTBNTNCRSCKRAVCTTYARNWHAKYSTWSAHSARCNVYCMGKRC 206

QY 305 TTGTCGGCTGTAAAGTACCGAGTACAGCTTCCTCGCTCGAACAGACTAACCGACCTC 364

DB 205 GCTHSSNCGCTHCRDRYSHRSMWBBANACNRBNANNSDSTNAGMNYNCKTBSGBAKHS 146

QY 365 CACACCC 371

DB 145 MMBSTCB 139

RESULT 7

AK119529/c

LOCUS

AK119529.1

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone:001-203-C05, full insert sequence.

ACCESSION

AK119529

VERSION

AK119529.1

KEYWORDS

FLI cDNA; oligo capping.

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team;

Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurotsaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

2

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Takami-Takeda, Y., Tagawa, A., Takahashi, F., Takaki-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice

Unpublished

3 (bases 1 to 1598)

Kikuchi, S.

Direct Submission

Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 32K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M., and Nakahama, Y.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,

Query Match	3.8%	Score 41.4;	DB 8;	Length 1618;
Best Local Similarity	53.4%;	Prod. No. 3.8;		
Matches	87;	Conservative	0; Mismatches	76; Indels
			0; Gaps	0;
QY	560	CCGTCCAACTTGGAGGCTATCTCAGAGCCGAGGAAGTTGGCGCAACTGGCAGCGTACG	619	
Ddb	1485	CCATPCGGGTTACAGATCGCTGTTGAAGATGGTACGTGGCGCAGGGCTGCGGCATGT	1426	
QY	620	TTGCCAGGAGGCCACGACGACGACGCGCGGCGGCACACCCCGACGCCCTCACTGCAA	679	

Db 1425 CGGCCACCGAGCCACCGCGGAGCAGCAGCAGCGCGCGCGCTTGAGCGCT 1366

Qy 680 CGAGCGCTCCGAACTTGAGCGGACACTTACCTTCCCTG 722

Db 1365 GCAGCGCCACCCACTTCCGCGTCCCTCTGCACTTCCGCTG 1323

RESULT 9

AK121636/c

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone:J033051N12, full

insert sequence.

AK121636

AK121636.1 GI:37991259

VERSION

FLI CDNA; CAP trapper.

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of

Agricultural Sciences Rice Full-length cDNA Project Team;

Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,

Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,

Kojima,K., Namiaki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,

Ohtsuki,K., Shishiki,T., Foundation of Advancement of International

Science Genome Sequencing & Analysis Group; Ootomo,Y., Murakami,K.,

Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,

Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Kie,Q., Lu,M.,

Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,

Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,

Kusumegi,T., Oka,Y., Ueda,M., Matsubara,K., RIKEN;

Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,

Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,

Kagawa,T., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,

Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,

Yoshino,M. and Hayashizaki,Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

REFERENCE

AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,

Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,

Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,

Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,

Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,

Kagawa,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,

Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,

Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,

Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,

Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakahama,Y.,

Nakamura,M., Namiaki,T., Narikawa,R., Niikura,J., Nishi,K.,

Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M.,

Ooka,H., Osato,N., Ota,Y., Ootomo,Y., Ryu,R., Satoh,H., Sakai,C.,

Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K.,

Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y.,

Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M.,

Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,

Tanaka,T., Tomaru,A., Tota,T., Tsunoda,Y., Ueda,M., Waki,K.,

Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J.,

Yokomizo,S. and Yoshimura,A.

Collection, mapping, and annotation of 28K full-length cDNA clones

from japonica rice

Unpublished

REFERENCE

AUTHORS

Kikuchi,S.

Direct Submission

Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of

Agricultural Sciences, Department of Molecular Genetics, Head of

Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki

305-8502, Japan (E-mail:skikuchi@nias.affrc.go.jp,

Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 32K full-length cDNA clones from japonica

rice.

URL : http://cdna01.dna.affrc.go.jp/cDNA/

NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,

Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,

Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiaki,T.,

Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T.,

Yamamoto,M. and Nakahama,Y.

FATS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y.,

Fujimura,T., Ikeda,R., Iehibiki,J., Kawamata,M.,

Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M.,

Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M.,

Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M.,

Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.

Genome Exploration Research Group in Riken Genomic Sciences Center

and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,

Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,

Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,

Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,

Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,

Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,

Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,

Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,

Ota,Y., Saitoh,H., Sakai,C., Sakai,K.,

Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K.,

Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M.,

Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,

Takaku-Akahira,S., Tanaka,T., Tomaru,A., Tota,T., Waki,K.,

Yasunishi,A. and Hayashizaki,Y.

Location/Qualifiers

1..1756

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="J033051N12"

FEATURES

source

Query Match 3.8%; Score 41.4; DB 8; Length 1756;

Best Local Similarity 53.4%; Pred. No. 3.9;

Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 560 CCCTCCAACTTGGAGGCTCTCTCTGAGCCGAGGAAGTTGGCGCAACTGGCAACGCTACG 619

Db 1485 CCATCCGGTTTCAGAGCTGGTTTGAAGATGTTGACGTGGCGAGCGCTGGCCCATGT 1426

Qy 620 TTGCCACGAGGCGCCACGACGCGCGCGAGCGCAACCCGACGCCCGCTCACTGCAA 679

Db 1425 CGGCCACCGAGCCCGCCGCGCGAGCGAGCGAGCGAGCGCGCGCGCTTGGCGCCT 1366

Qy 680 CCAGCGCTCCGAACTTGAAGCGGAAACACTTTTACCTTTCCCTG 722

Db 1365 GCAGCGCCACCCACTTCCGGTGCCTCTCTGCACTTCCGCTG 1323

RESULT 10

AX655393

LOCUS

AX655393

DEFINITION

Sequence 5263 from Patent WO03000898.

AX655393

ACCESSION

AX655393.1

VERSION

GI:29158207

KEYWORDS

ORyza sativa

SOURCE

ORyza sativa

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1

Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,

Karagiri,F., Quan,S., Tao,Y., Whitam,S., Xie,Z., Zhu,T. and Zou,G.

TITLE	Plant genes involved in defense against pathogens
JOURNAL	Patent: WO 0300089-A 5263 03-JAN-2003; Syngenta Participations AG (CH)
FEATURES	Location/Qualifiers
source	1..2000
ORIGIN	/organism="Oryza sativa" /mol_type="unassigned DNA" /db_xref="taxon:4530"
Query Match	3.8%; Score 41.4; DB 6; Length 2000;
Best Local Similarity	7.5%; Pred. No. 3.9; Indels 0; Gaps 0;
Matches 51;	Conservative 307; Mismatches 323;
QY	345 GAACAGACTAACCGGACCTCCACACCGCTTTTAAGCTCAGTACATAACGAAATCTCGTCGCAA 404
DB	104 GRRGMRSRWRWGRYRRCARSGRMAGGSGRMWGKSRMSYMMWCYARGCGSKRKSXG 163
QY	405 CGACAGCGGATGTTCTACGTAATTTGTTGGCTAGACGACACCAAGAAGAACCCATTGACGT 464
DB	164 GSWGKTCTRRGARGGSGSSGAKTKSGMSMKRMWMSGSGSGGSRSAYSRYTYGTSRYKGT 223
QY	465 CTTCGGATCAACTATCGGTGTCAATTTCCGGAACACCGCGCGGACTTCGCGGACTCTA 524
DB	224 YKMTTYSASCMRAYMTTYSWACSSYTWCRSKRSMWKRMRWRSRYSYGYWSYK 283
QY	525 TTCCAAGGCTTCGTTCGACCTTCGGAATTAACCTACCGTCCAACCTGAGGCTATCTCAG 584
DB	284 MMCTAYKSYSRWCYWRGGWGRATRYWGRGYMSRMAMMYKMYVRYGKMGKRW 343
QY	585 GACGAGAAAGTTGGCGCAACTGGCAAGCTACGTTGCGCAGGAGGCCACGACGACG 644
DB	344 AGRMWRSMCRWSKACYMRWRMTTRRRRWAUKSSRTSRKKRQKWRKRYKMRGY 403
QY	645 CGCCGAGCGACACCCGAGCGCGCTCACTGCAACACGCGCTCCGAACTTGAAGCGGA 704
DB	404 SRMSCKRARWKKRSGRAWKMGCGCGMTCRMKSYGMWRWSKRMASKYKWRMRTRW 463
QY	705 ACACTTTACCTTTCCCTGGCTAGAAAATGGCGTGGATCAATTACGACCGCACCGCGCAA 764
DB	464 RKKCSRRTTWGKTGGMMGTMGRCRYKKSNGMKRCRRRWGRWYMRWKRYSARYT 523
QY	765 CGAAATTTCAAAGTTACTGTCGCTTCGGGCAAGTACGCCCTACGCTAATTTGGGTAAAC 824
DB	524 MRYCARKKYSYSAARKARCWYRGKYYWAGMMKRYKRYMYKMMWYKRYKSKSYCK 583
QY	825 CGTGCTCGCTCGTAGCGCAACGATCGCGCTCTCATGTAAATTTCCATCGTCACCAG 884
DB	584 MSYASCMKSARKAGAKYCKRMSAWSMSRSRCKRCASKRSAKRYAMWGMSG 643
QY	885 AAACATGTGACCCCGCACCGAAATTAGACACGGTCTCGCAAGACGACGAAAGCTTC 944
DB	644 RMRWKSCTCYWRKWSKSTCTWYMYSKYTYAKYGSYWRYYRAWCMMWRWYTYR 703
QY	945 CCAAACTAGAGAGGAATCGGAAATTTGGACCCATGTTGGTGGTTCGGAATAAACAAGG 1004
DB	704 SYMTYMAWYTSRTMAWTGMKYSGRYWTWYKCKSKWYRSMWYWSWMAKTWMMWR 763
QY	1005 CGCTGACGAGATAGTGA 1025
DB	764 YATRMWMMWYRYSKMYTWCT 784
RESULT 11	
AP003261	
LOCUS	138025 bp DNA linear PLN 27-NOV-2003
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:p0471B04.
ACCESSION	AP003261 BA000010
VERSION	AP003261.3 GI:22202658
KEYWORDS	Oryza sativa (japonica cultivar-group)
SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	

/db_xref="GI:22202659"
/translation="MSPPSVICQFGDTTYYKUVUGLAWETOKETMRKYFQFGEHLE
AVITDKNTRSGKGVFTFRIDGRACRACVDFPVIDGRANCLNLSGVQSRPPT
POHGARSFRVMSQSQAQIGGLGAAPFPSPHATFHYAIPOGLFHYGVGSPYSPDY
SYNTYNYFGCAQYFYGAAAAAAGAMTGGSPVPYFQFQSGSTTWTNVSQO
GYNLOYPMFSTVASTAAATGFAQQVGGFSLAASPQAQAAEKANGKNNPQAL
TOGGFMILGANWTKVTSKQLIEQBCHFKISEXP"
complement(join(17199..17342,17914..18008,18141..18178,
18784..19073))
/genes="P0471B04.2"
complement(join(17199..17342,17914..18008,18141..18178,
18784..19073))
/gene="P0471B04.2"
/codon_start=1
/product="putative SnRK1-interacting protein 1"
/protein_id="BAC07318.1"
/db_xref="GI:22202660"
/translation="WASGALARLLIGRRAAATPILARPPAKARASRRPQBPAPSED
EDDFAGEVAAPFTEGISKPLAEVLRELGKRVPEALVKTRVEDGFSUKYIPFWHLVVK
NITLHAPESVGRSVISYSSDGKSVSYRVTLHGTDAIYREATGTSADDTGYGDP
VOKARMAFRACCTRLIGLHLHYHEDMS"
join(24593..24925,25029..25403)
/genes="P0471B04.3"
join(24593..24925,25029..25403)
/gene="P0471B04.3"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAC07319.1"
/db_xref="GI:22202661"
/translation="WARRQGVILRGRCRRRERGEREQVRLPLASTWTGRRGRG
GRRRGREGAGRTARQSGSPRRSSSSMTSRSEHQILSDQRTGLSNAWDS
TGLPVMSKVIQORGAVRRGRGRRGAGVAGEDGSGWGCCHGRGRSGA
GMAAADGADAAAEADGLGGCGRRFPERRAALLNCKSSRRSPFISTPLAAPS
PSAAILSSGAMLPNRR"
join(25836..26033,26372..26471,26764..26924,27200..27397,
27476..27568,27766..27888,28671..28918,29038..29209,
29441..29527,29557..29820,30449..30611,30873..30960,
31044..31113,31600..31728,32356..32445)
/genes="P0471B04.4"
join(25836..26033,26372..26471,26764..26924,27200..27397,
27476..27568,27766..27888,28671..28918,29038..29209,
29441..29527,29557..29820,30449..30611,30873..30960,
31044..31113,31600..31728,32356..32445)
/gene="P0471B04.4"
/notes="hypothetical protein
similar to Arabidopsis thaliana chromosome 5, At5g49880"
/codon_start=1
/protein_id="BAC07320.1"
/db_xref="GI:22202662"
/translation="WILRTPQKRBRADVDVDAATATTRSPVSDRLVLYDRPTAL
VFAGVPEPMDMVCYHCRQMVKSEFVVALDTAEKQVQYRATLDDMEELSKSEDE
RATCDKLNLYQELAAATKGRSAQERLLKEVGDFQRYCDQIKIGELSTQKKEI
DRIRAAESSASAKESVLEGNLQRLSENSEKALKKLSYLDQDTKLISKLNA
ELERMLRALNSEDRAKLNLEOLDKQOLDESRENEMEHRLNCSLSVETPSPD
DQKLIQLEELRYNEKEVDEARRLKSHTNVLLKEXILSEQGERAEMLSTQKE
ISAKQKLELEASTALLUSNPVDSFGDIPQKADIQKQALTNLKNVGEVLSQKE
LKVALEFADLSQRAEGEATLAKREASATREIKRELILAAISEERDRKRAVSK
SRGDADSSKCFATILNGMNIISCEVTIYMESDLSRMEKVTVELESTIRDOREL
QOHTLNLNMXELSTESKAKSLERDQOLRSQVALLLESKLGHGYSASSTKVLKRVN
TLAVNEAKQTLLEALQAKTKKERLQAVEELKGQDVTGVVDVNIAEKLAQKNQIA
TEKREERYKAVPARIISVFRKACSLFGYKIVMDQQQNGIPVTRILHSVVAQSD
DEKLEPFDESGSTNIIVDIFIRKMSIFAPFANLTMSFNKRSIC"
join(33470..33475,33573..33794,33969..34033,34129..34322,
34573..34839,34942..35029,35125..35250,35325..35434,
35557..35705)
/genes="P0471B04.5"
join(33470..33475,33573..33794,33969..34033,34129..34322,
34573..34839,34942..35029,35125..35250,35325..35434,
35557..35705)
/gene="P0471B04.5"
/notes="contains ESTs C91866(E31978),AUI66151(E31978)"
/codon_start=1
/product="putative elicitor response protein"

/protein_id="BAC07321.1"
/db_xref="GI:22202663"
/translation="MMTLTSERQPOEKPRPRPPLPKAVAALCVAFVVGILLIS
GRVVVLPFGSSPSASNYTSFSTGCENRAKLGESNPTDINNEVSRTHALQSLLDK
AVSLEMLAVERARSAACAGTAVGPKAFVVGINTAFSSKKRDSLRDWTW
PRGDKRLREKGIIVIRFVIGSGAAAGDPLDRAVDAEDAKDFLRLDHVEGVH
ELSKRYVTYTTAVATWADPYVKYDDDVHVNLMGLTSRLAKYRTRPRVVGCMKSGP
VLQSGVKYHEPEYKFGDEGNKIFRHATGIIYAVSADLAAYISINQPIILHRFANEDY
SLGAWLGLGVEHVDVDRSMCCATPPDCEWKKRAAGNVCVASFDMSCSGVCKSVDRMKHI
HRACGEGQAGVWSVAT"
join(36438..36570,36637..36956)
/genes="P0471B04.6"
join(36438..36570,36637..36956)
/gene="P0471B04.6"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAC07322.1"
/db_xref="GI:22202664"
/translation="MSLTFSVLVWPSRGRPRPQVRGRGSSSSVGGGGBESHWTG
PATERRVRRRRRRRRRTVLPNNIGACPLSPRCGADGATPRSGRGFPDAP
TTAPDPLRVTVRDGEVSDGARQQRQWMDGQRGREREREAATADK"
complement(join(38223..38304,39187..39287))
/genes="P0471B04.7"
complement(join(38223..38304,39187..39287))
/gene="P0471B04.7"
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAC07323.1"
/db_xref="GI:22202665"
/translation="MVNDRKSVTGLPVTCDWTDNNAPTSHLLPAEWSTFSHSIDDT
CKVIRKIIIVSYSAIG"
join(43057..43393,43868..43997,44238..44453,44554..45171,
45260..46109)
/genes="P0471B04.8"
join(43057..43393,43868..43997,44238..44453,44554..45171,
45260..46109)
/gene="P0471B04.8"
/notes="contains ESTs AU057294(S21318),AU057295(S21318)"
/codon_start=1
/product="putative male sterility 1 protein"
/protein_id="BAC07324.1"
/db_xref="GI:22202666"
/translation="MVVNSPLKRRTRVEARDFAGPPAGDGAAGTFFREAVRGFLA
RYARLLPLPSIFSPAAAAAPPHTLWRSVRGVEGDEGGGAGVAVLVVEEDVLRGR
SVYCDQCRVVGNSGHPVCGKRYHFIENDNNQVCGKRHSCLSCGTPVAGESRCLLC
NFDMDGELEECGTMHLDNTHLHVAHVHANGHLLRVNREGSGRCLTGRDIMSEFW
Query Match 3.8%; Score 41.4; DB 8; Length 138025;
Best Local Similarity 53.4%; Pred. No. 7.5;
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
Qy 560 CCGTCCAACTTGAGGCTTATCTCAGGACCGAGGAAATTTGGCGCACTGGCAAGCGTACG 619
Db 95038 CCATCCGGTTTACAGCTGCGTTTGAAGATGTTGAGTGGCGCAGCGCTGCGCCATGT 95097
Qy 620 TTGCCAGGAGCCACGACGACGCGCGGAGGAGCAACCCGACCGCTCTACTGCAA 679
Db 95098 CGGCGACCGAGCCACCGCGGAGCGAGCAGCAGCAGCGCGCGCGCGCGCTTGGAGCGCT 95157
Qy 680 CCAGCGCTCCGACTTGAAGCGCAACACTTACCTTTCCTG 722
Db 95158 GCAGCGCACCCACTTCCGCGTGCCTCTCTGCACCTTCGCGCTG 95200
RESULT 12
AP003227
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
PAC clone: P0018C10.
ACCESSION AP003227 BA000010
VERSION AP003227.3 GI:21952783
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM

Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartideae; Oryzeae; *Oryza*.

REFERENCE

AUTHORS

1
Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T.,
Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y.,
Antonio, B. A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K.,
Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K.,
Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y.,
Idonuma, A., Iijima, M., Ikeda, M., Ikeda, M., Ikeno, M., Itoh, S., Itoh, T.,
Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S.,
Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T.,
Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M.,
Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M.,
Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T.,
Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K.,
Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K.,
Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J. H., Kim, H. I., Eun, M. Y.,
Yano, M., Jiang, J. and Gojobori, T.
The genome sequence and structure of rice chromosome 1

Nature 420 (6913), 312-316 (2002)

TITLE

JOURNAL

MEDLINE

PUBMED

2237376

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

2 (bases 1 to 200193)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
1-1-2, Tsukuba, Ibaraki, 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Jul 24, 2002 this sequence version replaced gi:21202836.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative', and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
This sequence of P0018C10 clone has an overlap with P0471804
clone (DBJ: AP003261) at the position 1 to 89,225 of 5' end and an
overlap with B1065810 clone (DBJ: AP003561) at the position
158,410 to 200,183 of 3' end. Detailed information on overlap
assembly quality together with annotation of this entry is
available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES

source

1. 200183
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clones="P0018C10"
join(6052..6159,6257..6388,6483..6612,6719..6805,
7232..7302)
/gene="P0018C10.1"
join(6052..6159,6257..6388,6483..6612,6719..6805,
7232..7302)
/gene="P0018C10.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC06200.1"

gene

CDS

/db_xref="GI:21952784"
/translation="MWTSLIARSTTCLQTANDTTTAPAKPAGAGPOLKLLADADISLE
EFGSLCPPICTGELLITPDSNFAVVTREPLLYVQVTRLCRCGFFVSTOICNNLVDA
AGTQFLQAVGELVEGRAPAGDELVHRRFLSGPDDVAALRDQLPTRLSPGCHVSKTK
QKYCQETFGSEFV"
complement(join(14216..14383,14464..14772))
/gene="P0018C10.2"
complement(join(14216..14383,14464..14772))
/gene="P0018C10.2"
/note="contains EST AU058082(B30867)
unknown protein"
/codon_start=1
/protein_id="BAC06201.1"
/db_xref="GI:21952785"
/translation="MGRSRPWFQDPSAPPQPVVAAAADAGEASTSLRDFGTSMD
AISFGFAATAILVSMFLMIAEHLIKPHVFPPLASALRPARRHGVSPAGKLRSP
MVETVQAADLSVLMPCQRYPTTIAQAPLPAPCPREGVHPPHDDHVSHTNMP"
complement(join(21230..21448,21542..21620,22147..22244))
/gene="P0018C10.3"
complement(join(21230..21448,21542..21620,22147..22244))
/gene="P0018C10.3"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC06202.1"
/db_xref="GI:21952786"
/translation="MGYGRILSLRSLVTSLNQLNPTDNLHQNNTKHADVETDDLP
KNKGALSNVSAEKVEAKKANKGRRGSGYSIRPPVHGYAASDRSSNPLGDL
DFLLTMGERAVHLPAKPRPVKAASIRSP"
join(26342..28883,29797..30149)
/gene="P0018C10.4"
join(26342..28883,29797..30149)
/note="contains ESTs
C96750(C10643), C22379(C12740), C22380(C12740)"
/codon_start=1
/product="putative receptor protein kinase"
/protein_id="BAC06203.1"
/db_xref="GI:21952787"
/translation="MLPQQLQIYLCFILLSLKFGISASLPLETALLDIKSHLEDDPN
YLGWDSHSPCQFYGTQCTGGVIGISLNASLSGTISSEFSLQSQRTLELGN
SISGTIPAAALNCTNQLVNLSTLQKLPDLSTFINLOVLDTSTNFFSPFPFVWG
SISGTITELGALGNNFEGDVPESIGKLNLTWFLGQCNLRGELPVSIFDLVSLGTLID
FSRQIIGVPIALSNRLNWKILYQNLITGEIPPELAHLTLISEPVSQNLQSLGIL
PEKIANLKUKIFHIYRNNFSGVLPEGLDEFESFSTYENQSGKFPANLRSPSL
NAIDISNYSFGEPRFLCQNNKQLFLALDNNFSGFPSPSSCKTLQRFRIQSNQ
TGRTHSGTWLPNAVIVDANKFVGISDGIASLNOLYVHNNVSGELPMELGK
LSLQKLVANNRESGQIPAQIGSLKOLSLHLQNALGSI PPDIGMCSLVDLNLA
LMTAGDDAFSENDGLCIAGVSEGRQNAIRYC PNWDNNHQNFSQRFLFVLLIVTS
LVLLSLGLACLYENYKLEQFHSKGDIESGDDSDSKVLESFPELDPESICNDVD
NLICGGTGKVVYVNGLYDAIRREFKAGQPELDWEKRYIAVGTAKGIMYLHDCS
LTGSESLVYVYVNGLYDAIRREFKAGQPELDWEKRYIAVGTAKGIMYLHDCS
PAIIRDIKSTNILLDEYEAKLADFGIAKLVEGSLSPCFAGTHGYNAPELAYSILVT
EKSDVYSGIVLLELTGRSPKDDQDGLDIVSWSSHANQNPAAVLDPKVSASHAS
EDMTKLVNAILCTVLPSPRTWREYVLMILDISANGAKNNKDKK"
complement(join(30791..31219,31321..31545,31728..32088,
32199..32292,32370..32603,33125..33179))
/gene="P0018C10.5"
complement(join(30791..31219,31321..31545,31728..32088,
32199..32292,32370..32603,33125..33179))
/gene="P0018C10.5"
/codon_start=1
/product="putative amino acid permease 6"
/protein_id="BAC06204.1"
/db_xref="GI:21952788"
/translation="MNKNAAPEDVESGEHRTGTWTATAHIIVATIGSGVLANSWS
AQLGWAGPLALAGVCTITATISWAIIRSDCIHRNGAAGAAACDNTSATVLMFAFS
MCGTAQVYVNLGTMVGTITATISWAIIRSDCIHRNGAAGAAACDNTSATVLMFAFS
IVQVLSQFPGLIEHTVYVAAVMSFAYSPFISGLVSAEVSCHGSLRIGQATGTA
SSSKIMNVLLALGNIAFYFAEVLIEIQTLPKSPENKMKAAWYGATTIFY
ISVGCAGYAAFGSDAPGNIPASGFPFLVDIANMCLILHLIGIYQVYQIPATME
RWISRWPEAKFINSEYTVNVPLIQSGSVTVAPKLVLTVTVATTVAAMIFFNA
VLGLLGAFSFWPLTVYFIPSMHIAQEKITRGRWYLLQGLSMVCLMISVAVIGSVTD


```
Db 71347 GACCATCT 71339

RESULT 14
AY135117
LOCUS
DEFINITION
Drosophila melanogaster isolate 1 corkscrew phosphatase splice
variant (csw) gene, alternatively spliced, complete cds.
ACCESSION
AY135117
VERSION
AY135117.1 GI:23344935
KEYWORDS
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 5513)
REFERENCE
Riley,R.M., Jin,W. and Gibson,G.
AUTHORS
Riley,R.M., Jin,W. and Gibson,G.
TITLE
Contrasting selection pressures on components of the Ras-mediated
signal transduction pathway in Drosophila
JOURNAL
Mol. Ecol. 12 (5), 1315-1323 (2003)
MEDLINE
22582092
PUBMED
12694293
REFERENCE
2 (bases 1 to 5513)
AUTHORS
Riley,R.M., Jin,W. and Gibson,G.
TITLE
Direct Submission
JOURNAL
Submitted (23-JUL-2002) Department of Genetics, North Carolina
State University, Gardner Hall 7614, Raleigh, NC 27695, USA
FEATURES
Location/Qualifiers
1..5513
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/isolate="1"
/db_xref="taxon:7227"
/country="USA: Ann Arbor, MI"
/ncbi="genotype: isofemale"
<161..>5449
/gene="csw"
/join(<161..174,757..944,1017..1143,2096..2258,2951..3130,
3199..3316,3445..4382,4461..4551,4666..4855,4921..5449)
/product="corkscrew phosphatase splice variant A"
/join(161..174,757..944,1017..1143,2096..2258,2951..3130,
3199..3316,3445..4382,4461..4551,4666..4855,4921..5449)
/gene="csw"
/notes="CswB"
/codon_start=1
/product="corkscrew phosphatase splice variant A"
/protein_id="AA017607.1"
/db_xref="GI:23344936"
/translation="MSSRRWFPTTISGIEAKLQEQFGDGSFLARLSSNPGFTLS
VRGNEVTHIKIONGDPPDLYGKGFATLPBVQYWMENGELKNGQALTELKQPLI
CAEPTTERFHNLGSKAEKILLRKNGSFLVRESQKPGDFLVSVRTDDKVTHVM
IRWQKIVDGGESFGTSELIDYKKNPMVETCTVHLRQPNATRIITAAGINAR
VEQLVKGWVEFEFLQSDRTFRNKGYSKQENLKNRYLPIYDTRVRLGLDVEH
SVAGAEINANTIRLPTDGLYMSSESLSNVPSCPAQTQRCNSCQLOK
TCVQCAVKAILPYSNCATCSRKSDLSLKHKSSESASSPSGSGSGSGTSGVS
SVNGPCTNLSTGTAGCLVGLLRHNSDGAISMAEREREREMFTYIATQG
CLLTQVNTVFWNNWQENTRVYMTKEYERKEKCARVWPDGSEQFGHARQ
CVSNTSYTLRFLVSWRDPARFIHYFWPDHGVDPADPCVNLFPQVNTRO
SHLAQAGEKPGICVHCASGIGTGTIVDMLDQIVRNGLDITDQRTQWVRSQ
RSLGVAEQYKFFVAVQHYQTLIARAEQSLQVREYTNKYTGEGINDSOR
PLPAPTSSILVPSKPTLPTSDTLGTMGVMGVGNKHSKQPPPLPVNCCN
NNGNIGNSCSGSGSTSSNGSNNGINALLGIGLGGNNRKNFYSLSLKQ
QKREEQAPAGAGWQAPPLPRPGILKLTSPVIFQNSKTFPKT"
/join(<1332..1960,2096..2258,2951..3130,3199..3316,
3445..4382,4461..4551,4666..4855,4921..5449)
/gene="csw"
/product="corkscrew phosphatase splice variant B"
/join(1332..1960,2096..2258,2951..3130,3199..3316,
3445..4382,4461..4551,4666..4855,4921..5449)
/gene="csw"

gene
mRNA
CDS
RESULT 15
AY135123
LOCUS
DEFINITION
Drosophila melanogaster isolate Reids2 corkscrew phosphatase splice
variant (csw) gene, alternatively spliced, complete cds.
ACCESSION
AY135123
VERSION
AY135123.1 GI:23344953
KEYWORDS
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 5514)
REFERENCE
Riley,R.M., Jin,W. and Gibson,G.
AUTHORS
Riley,R.M., Jin,W. and Gibson,G.
TITLE
Contrasting selection pressures on components of the Ras-mediated
signal transduction pathway in Drosophila
JOURNAL
Mol. Ecol. 12 (5), 1315-1323 (2003)
MEDLINE
22582092
PUBMED
12694293
REFERENCE
2 (bases 1 to 5514)
AUTHORS
Riley,R.M., Jin,W. and Gibson,G.
TITLE
Direct Submission
JOURNAL
Submitted (23-JUL-2002) Department of Genetics, North Carolina
State University, Gardner Hall 7614, Raleigh, NC 27695, USA
FEATURES
Location/Qualifiers
1..5514
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/isolate="Reids2"
/db_xref="taxon:7227"
/country="Portugal: Mediera"
/notes="CswB"
/codon_start=1
/product="corkscrew phosphatase splice variant B"
/protein_id="AA017608.1"
/db_xref="GI:23344937"
/translation="MLFNKLEKLSLSSIGNVNHVKLOEQVYNNNNNNNNNNNNNNNN
NAYNORNEFEYRAIOAHYGSKRSEERSGKFKASKGRKAKVTPPTETPAOBPA
CKMCMTHDELAQILKGVAGDAQRNRRLORRRPLSAQPSAASASTSTSLHRL
TPSPQASYPATPTSTATPPAPAFGGASCNSTLSLATMRVQLHGTWTHGNLJSG
KEAKLILRKNKNSFLVRESQKPGDFLVSVRTDDKVTHVMIRWQKIVDGGESF
GTLSELIDYKKNPMVETCTVHLRQPNATRIITAAGINARVEQLVKGWVEFEFL
QODSDRTFRNKGYSKQENLKNRYLPIYDTRVRLGLDVEHSAEYINANYIRLP
TGDDLYNMSSESLSNVPSCPAQTQRCNSCQLOKTCVQCAVKAILPYSN
CATCSRKSDLSLKHKSSESASSPSGSGSGSGTSGSGVNGPCTNLPTLSTGA
GCLVGLLRHNSDGAISMAEREREREMFTYIATQCLLTQVNTVDFWNN
VMQENTRVYMTKEYERKEKCARVWPDGSEQFGHARQCVSNTSYTLRFL
VSWRDPARFIHYFWPDHGVDPADPCVNLFPQVNTROSHLAQAGEKPGICVH
CSAGIGTGTIVDMLDQIVRNGLDITDQRTQWVRSORSLQVREYTNKYTGEGIND
AVUHYQTLIARAEQSLQVREYTNKYTGEGINDSORPLPAPISISILVPSKT
PUTPTSDTLGTMGVMGVGNKHSKQPPPLPVNCCNNGNIGNSCSGSGGS
STSSNGSNNGINALLGIGLGGNNRKNFYSLSLKQKQKREEQAPAGAGWQ
PAPPLPRPGILKLTSPVIFQNSKTFPKT"
ORIGIN
Query Match 3.6%; Score 39.6; DB 3; Length 5513;
Best Local Similarity 52.4%; Pred. No. 15;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 569 TTGAGGCTTCTCAGACCGAGGAAAGTTGGCGCACTGGCAAGCGTACGTTGCCACGG 628
Db 3979 TTAAGACCTACATCGCCACCCAGGGGTGTCTGCTCACCCAGCAAGTGAACACGCTACGG 4038
Qy 629 AGGCACGACGACGACGCGCGGCGACAAACCCCGCGCCGCTCACTGCAACACGCGCCT 688
Db 4039 ACTTTGGAACNTGGTCTGCGCAGAGAACACGCGGGTGTATCGTCATGACACCAAGAGT 4098
Qy 689 CCGAAGTTGAAGCGGCAACACTTTACCTTCCCTCCCTGGCTAGAAAATGG 734
Db 4099 ACGAGCGCGCAAGAAAGTGGCGCCGCTACTTGGCCAGCAGAGGG 4144
RESULT 15
AY135123
LOCUS
DEFINITION
Drosophila melanogaster isolate Reids2 corkscrew phosphatase splice
variant (csw) gene, alternatively spliced, complete cds.
ACCESSION
AY135123
VERSION
AY135123.1 GI:23344953
KEYWORDS
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 5514)
REFERENCE
Riley,R.M., Jin,W. and Gibson,G.
AUTHORS
Riley,R.M., Jin,W. and Gibson,G.
TITLE
Contrasting selection pressures on components of the Ras-mediated
signal transduction pathway in Drosophila
JOURNAL
Mol. Ecol. 12 (5), 1315-1323 (2003)
MEDLINE
22582092
PUBMED
12694293
REFERENCE
2 (bases 1 to 5514)
AUTHORS
Riley,R.M., Jin,W. and Gibson,G.
TITLE
Direct Submission
JOURNAL
Submitted (23-JUL-2002) Department of Genetics, North Carolina
State University, Gardner Hall 7614, Raleigh, NC 27695, USA
FEATURES
Location/Qualifiers
1..5514
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/isolate="Reids2"
/db_xref="taxon:7227"
/country="Portugal: Mediera"
/notes="CswB"
/codon_start=1
/product="corkscrew phosphatase splice variant B"
/protein_id="AA017608.1"
/db_xref="GI:23344936"
/translation="MSSRRWFPTTISGIEAKLQEQFGDGSFLARLSSNPGFTLS
VRGNEVTHIKIONGDPPDLYGKGFATLPBVQYWMENGELKNGQALTELKQPLI
CAEPTTERFHNLGSKAEKILLRKNGSFLVRESQKPGDFLVSVRTDDKVTHVM
IRWQKIVDGGESFGTSELIDYKKNPMVETCTVHLRQPNATRIITAAGINAR
VEQLVKGWVEFEFLQSDRTFRNKGYSKQENLKNRYLPIYDTRVRLGLDVEH
SVAGAEINANTIRLPTDGLYMSSESLSNVPSCPAQTQRCNSCQLOK
TCVQCAVKAILPYSNCATCSRKSDLSLKHKSSESASSPSGSGSGSGTSGVS
SVNGPCTNLSTGTAGCLVGLLRHNSDGAISMAEREREREMFTYIATQG
CLLTQVNTVFWNNWQENTRVYMTKEYERKEKCARVWPDGSEQFGHARQ
CVSNTSYTLRFLVSWRDPARFIHYFWPDHGVDPADPCVNLFPQVNTRO
SHLAQAGEKPGICVHCASGIGTGTIVDMLDQIVRNGLDITDQRTQWVRSQ
RSLGVAEQYKFFVAVQHYQTLIARAEQSLQVREYTNKYTGEGINDSOR
PLPAPTSSILVPSKPTLPTSDTLGTMGVMGVGNKHSKQPPPLPVNCCN
NNGNIGNSCSGSGSTSSNGSNNGINALLGIGLGGNNRKNFYSLSLKQ
QKREEQAPAGAGWQAPPLPRPGILKLTSPVIFQNSKTFPKT"
/join(<1332..1960,2096..2258,2951..3130,3199..3316,
3445..4382,4461..4551,4666..4855,4921..5449)
/gene="csw"
/product="corkscrew phosphatase splice variant B"
/join(1332..1960,2096..2258,2951..3130,3199..3316,
3445..4382,4461..4551,4666..4855,4921..5449)
/gene="csw"
```

```
gene
/notes="genotype: isofemale"
<161..>.5450
/genes="csw"
join(<161..174,757..944,1017..1143,2097..2259,2952..3131,
3200..3317,3446..4383,4462..4552,4667..4856,4922..5450)
/genes="csw"
/product="corkscrew phosphatase splice variant A"
join(161..174,757..944,1017..1143,2097..2259,2952..3131,
3200..3317,3446..4383,4462..4552,4667..4856,4922..5450)
/genes="csw"
/codon_start=1
/product="corkscrew phosphatase splice variant A"
/protein_id="AAN17619.1"
/db_xref="GI:23344954"
/translation="MSRRWFHTISGIEAEKLBQPGDGSFLARLSSNPGCATLS
VRGNEVTHIKIQNGDFDLVYGGKEFATLPELVYIMENGELKKNQAIELKQPLI
CAEPTTERFHGSLGKEAEKILERKNGSLVRESQKPGDFVLSVTRDDKTVTHM
IRWQKDYDVGGSFTLSLIDHYKRNPMVETCGTVVHLRQPNFATRIITAAGINAR
VEOLVKGFWEEFESLQDSRDTFSENEGYKOENLKNEYRNILPVDRVRLDVEH
SVAGAYIRNRYLFTDGLDLYMSSSESLSSVSPCPACTAATQRCNSCQLOKQ
TCQCAVKGAILPYSCATCSRKSDLSLKHKSSESASSPSSGSGPGSGTSGVS
SVNPGPTNLTGSLAGCLVGLLKRHSNDSGAVSISMAEREREREMFKTYIATQ
CLLTQOQNTVTFDNNVWQENTRVIMTKYERKKEKARYWDEGRSEQFQHARIO
CVSENSTDYTLREELVSWEDQPARIEHYHFQWPDHGVADPCVLNFDLVNTRQ
SHLAQAGEPGPICVCSAGIGRTGTFVIDMLDOIVRGLDTEIDITQRTQMVRSQ
RSLVQTEAQYKEVYAVQHYIQTLIARKRAEQSLQVGREYTNIKYIGIENDSQRS
PLPAISSLSLPSKTPLTPTSDGLTGLMGVGMVGNKSHKSKQPPLEPVNCCN
NNGNIGSSCSNGSGSTSSSSNGNINALLGIGLGLGNMRKSNFYSDSLKQQ
QOREQAPAGAKWQOPAPPLRPRPGILKLTSPVIFQNSKTFPKT"
join(<1332..1960,2097..2259,2952..3131,3200..3317,
3446..4383,4462..4552,4667..4856,4922..5450)
/genes="csw"
/product="corkscrew phosphatase splice variant B"
join(1332..1960,2097..2259,2952..3131,3200..3317,
3446..4383,4462..4552,4667..4856,4922..5450)
/genes="csw"
/codon_start=1
/product="corkscrew phosphatase splice variant B"
/protein_id="AAN17620.1"
/db_xref="GI:23344955"
/translation="MLFNKLEKLSLGLNVNHLKQEKQVYNNNNNNNNNNNNLNN
NATYNNQNEYERAIQAHYKSGRRSEERERKGFASKGRKAKVTPPTPEAOEPA
CKNCMTDELQAIKIGVAGADARNRDLQRRRPLSAQPSAASASTESLHRL
TPSPQASYATPPTPPQAPFAGSCSNLSLILATMRVQLHGYTFWPHGLSG
KEAKLILERKNGSLVRESQKPGDFVLSVRTDDKTVHMRQDKKYDVGGGESP
GTLSELIDHYKRNPMVETCGTVVHLRQPNFATRIITAAGINARVEQLVKGFWEEFESL
QQSRDTFGRNEGYKQENLKNEYRNILPYDHRVKLLDVHVSAGASTINAYIRLP
TDGDLNMSSESLSSVSPCPACTAATQRCNSCQLOKQTCVCAVKSAILPYSN
CATCSRKSDLSLKHKSSESASSPSSGSGPGSGTSGVSVNGPPTNLTGSGTA
GCLVGLLKRHSNDSGAVSISMAEREREREMFKTYIATQGLLTQOQNTVTFDNN
VMQENTRVIMTKYERKKEKARYWDEGRSEQFQHARIOCVSENSTDYTLREEL
VSWRQPARRIEYHFQWPDHGVADPCVLNFDLVNTRQSHLAQAGEKPGPICVH
CSAGIGRTGTFVIDMLDOIVRGLDTEIDITQRTQMVRSLVQTEAQYKEVY
AVQHYIQTLIARKRAEQSLQVGREYTNIKYIGIENDSQRSPLPAISSLSLPSK
PLTPTSDGLTGLMGVGMVGNKSHKSKQPPLEPVNCCNNGNIGSSCSNGSGS
STSSNGSGSNGINALLGIGLGLGNMRKSNFYSDSLKQQQREEQAPAGAKWQQ
PAPLPRPRGILKLTSPVIFQNSKTFPKT"
```

ORIGIN

```
Query Match          3.6%; Score 39.6; DB 3; Length 5514;
Best Local Similarity 52.4%; Pred. No. 15;
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 569 TTGAGGCCTATCTCAGGACCGAGAAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGG 628
Db 3980 TTAAGACCTACATGCCACCGACCGGCTGTCTGTCTACCCAGCAAGTGAACACGGTGACGG 4039

Qy 629 AGGCCAGCAGCAGCGCGGCGGAGCAACCCGAGCGCCGCTCACTGCAACGAGCGCT 688
Db 4040 ACTTCTGGAACATGCTTGGCAGGAGAACACGCGGGGTGATCGTCATGACCAACAGGAGT 4099

Qy 689 CCGAACTTGAAGCGGACACTTTACCTTTCCCTGGCTAGAAATGG 734
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 20:07:34 ; Search time 552 Seconds
(without alignments)
8357.868 Million cell updates/sec

Title: US-09-993-777-1_COPY_9874_10959
Perfect score: 1086
Sequence: 1 ATGCATCGCTACTTGAAC.....CGCCGACTCATAAAATG 1086

Scoring table: GAPop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

ALIGNMENTS

RESULT 1									
AAT33504									
ID	AAT33504 standard; DNA; 13473 BP.								
XX	AAT33504;								
DT	16-OCT-2003 (revised)								
DT	18-NOV-1996 (first entry)								
XX	Infectious laryngotracheitis virus short region sequence.								
XX	Infectious laryngotracheitis virus; ILTV; herpesvirus; attenuation;								
KW	vector; vaccine; chicken; poultry; immunisation; ds.								
XX	Gallid herpesvirus 1.								
XX	Key	Location/Qualifiers							
FT	repeat_region	1..273							
FT		/tag= a	"internal repeat region"						
FT	misc_feature	274..13371							
FT		/function=	"unique short sequence of ILTV"						
FT		/tag= b							
FT	CDS	complement(281..970)							
FT		/tag= c							
FT		/label= US2_gene							
FT	CDS	1059..2489							
FT		/tag= d							
FT		/label= Protein-kinase_gene							
FT	CDS	2575..4107							
FT		/tag= e							
FT		/label= UL47-like_gene							
FT	CDS	4113..4445							
FT		/tag= f							
FT		/label= ORF4							
FT	CDS	complement(4139..4519)							
FT		/tag= g							
FT		/label= ORF4_reverse_complement							
FT	misc_difference	4535							
FT		/tag= h							
FT		/note= "base 4535 is given as 's' in the specification"							
FT		4609..5487							
FT	CDS	/tag= i							
FT		/label= gg_gene							
FT		5697..8654							
FT	CDS	/tag= j							

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1086	100.0	13473	2 AAT33504
2	1086	100.0	13473	2 AAT33504 Infectiou
3	1084.4	99.9	3605	6 ABK90557 Infectiou
4	1084.4	99.9	18912	2 AAT44385 Infectiou
5	343	31.6	534	2 AAT33513 Infectiou
6	47.8	4.4	2000	7 ADA71938 Rice gene
7	41.4	3.8	2000	7 ADA71938 Rice gene
8	38.2	3.5	777	7 ACA38608
9	38.2	3.5	780	7 ACA40341
10	38.2	3.5	110000	4 AAI99682_06
11	38.2	3.5	110000	4 AAI99683_06
12	38	3.5	2526	5 AAF26455
13	38	3.5	2677	4 AB106927
14	38	3.5	2838	4 AB118521
15	38	3.5	4528	4 AB121135
16	38	3.5	10587	4 AB118520
17	38	3.5	19468	4 AB106926
18	38	3.5	21399	4 AB121134
19	37.6	3.5	813	6 ABQ13969
20	37.6	3.5	813	6 ABQ13968
21	37.6	3.5	6310	6 ABK31365
22	37.6	3.5	6310	6 AB170322
23	37.6	3.5	6310	6 AAS61269

/label= g60 gene
complement(5948. .7826)
/*tag= k

/label= ORF6_reverse_complement
8462..9766
/*tag= i

/label= gd_gene
9874..10962
/*tag= m

/label= GI Gene
complement(10617..11150)
/*tag= n

/label= ORF8_reverse_complement
11159..12658
/*tag= o

/label= gE gene
12685..13447
/*tag= p

/label= ORF10
misc_difference 13002..13003
/*tag= q

/note= "bases 13002-13303 are given as 'ss' in the specification"

/repeat_region
13372..13473
/*tag= r

/function= "terminal repeat region"

W09508622-A1.

30-MAR-1995.

16-SEP-1994; 94WO-US010628.

24-SEP-1993; 93US-00126597.

(SYTR) SYNTRO CORP.

Wild MA, Cochran MD;

WPI: 1995-139591/18.

P-PSDB; AAW00630, AAW00631, AAW00632, AAW00633, AAW00634, AAW00635,
AAW00636, AAW00637, AAW00638, AAW00639, AAW00640, AAW00641, AAW00642.

Recombinant attenuated infectious laryngotracheitis virus - for use in vaccines to protect poultry from infection from the virus, also methods of distinguishing between vaccinated and naturally infected birds.

Example 1; Page 79-94; 17pp; English.

The unique short region (AAT33504) of infectious laryngotracheitis virus (ILTV) genomic DNA contains genes (see also AAT33505 and AAT33510-13) that are associated with ILTV virulence. A deletion in those genes, esp. the glycoprotein-gI gene, glycoprotein gI gene, thymidine kinase gene, US2 gene, UL47-like gene or the glycoprotein g60 gene, will attenuate the ILTV. A gene for a foreign antigen may be inserted into the US2, UL47-like, ORF4, gG, g60 or gI gene to produce a recombinant ILTV. Recombinant ILTV may be used as a multivalent vaccine, esp. for use in poultry.
(updated on 16-OCT-2003 to standardise OS field)

Sequence 13473 BP; 3390 A; 3582 C; 3547 G; 2951 T; 0 U; 3 Other;

Query Match 100.0%; Score 1086; DB 2; Length 13473;
Best Local Similarity 100.0%; Pred.No.0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCATCGTACTTGGAACCTCGCTCTCCTTCGCCGCAGCGCTCGCACCCCTTCGCCGCG 60
9874 ATGGCATCGTACTTGGAACCTCGCTCTCCTTCGCCGCAGCGCTCGCACCCCTTCGCCGCG 9933

61 ATGGCAATCGTGATCATTGGAACAATGATGTTCGCCGAGGATTTGACGACGATGATCATCTG 120
9934 ATGGCAATCGTGATCATTGGAACAATGATGTTCGCCGAGGATTTGACGACGATGATCATCTG 9993

[illegible]

PT Recombinant infectious laryngotracheitis virus with deletion in the
PT glyco:protein G, g1 or US2 gene, etc. - useful for vaccines against
PT infectious laryngotracheitis in poultry.
XX Example 11; Page 88-103; 216pp; English.
PS
XX The nucleotide sequence of 13,473 bp of contiguous DNA (AAR44384) from
CC the unique short region of infectious laryngotracheitis virus (ILTV)
CC contains the entire 13,098 bp unique short region and includes 13 open
CC reading frames that encode proteins (AAR06782-94) of over 100 amino
CC acids, 8 of which show significant homology to other virus genes. Novel
CC recombinant, attenuated ILTV comprises the ILTV genome contg. a deletion
CC in the unique short region, esp. in the glycoprotein gg, g1, US2, ORF4,
CC UL47-like or g60 gene. The attenuated virus is useful as a vaccine
CC against ILTV. A foreign gene encoding an antigen e.g. from another avian
CC virus can be inserted into the US2, UL47-like, ORF4, g1, g60 or g1 gene
CC to provide a multivalent vaccine for chickens and other poultry. Deletion
CC of the gg or g1 gene provides as a negative marker to distinguish
CC vaccinated from infected animals. (Updated on 16-OCT-2003 to standardise
CC OS field)
XX
SQ Sequence 13473 BP; 3390 A; 3581 C; 3547 G; 2952 T; 0 U; 3 Other;
Query Match 100.0%; Score 1086; DB 2; Length 13473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGATCGCTACTTGGAACTCGCTCTCTTGGCGGAGCGCTCGACCCCTTCGGGCG 60
Db 9874 ATGGGATCGCTACTTGGAACTCGCTCTCTTGGCGGAGCGCTCGACCCCTTCGGGCG 9933
QY 61 ATGGGATCGCTACTTGGAACTCGCTCTCTTGGCGGAGCGCTCGACCGATGACATCGTG 120
Db 9934 ATGGGATCGCTACTTGGAACTCGCTCTCTTGGCGGAGCGCTCGACCGATGACATCGTG 9993
QY 121 ATCGTGGCGCTCGCGCCGAGCTCAATTCAGTGCAGCTATTTTCATGCTGGCCAG 180
Db 9994 ATCGTGGCGCTCGCGCCGAGCTCAATTCAGTGCAGCTATTTTCATGCTGGCCAG 10053
QY 181 AGACCCCAACAAACCTTACTCAGAAACCGTTCGGGTGCGGTTTCGGTCTGATATACAAAC 240
Db 10054 AGACCCCAACAAACCTTACTCAGAAACCGTTCGGGTGCGGTTTCGGTCTGATATACAAAC 10113
QY 241 CAGTGTACAGGAATAGCAGGAGCGCTTTGAATGACATCAGTGCAGTCTCTCT 300
Db 10114 CAGTGTACAGGAATAGCAGGAGCGCTTTGAATGACATCAGTGCAGTCTCTCT 10173
QY 301 GTTTTGTGGCTGTAACTGACGAGTACAGTCTCGCGCTCGAAACAGACTCAACCGGA 360
Db 10174 GTTTTGTGGCTGTAACTGACGAGTACAGTCTCGCGCTCGAAACAGACTCAACCGGA 10233
QY 361 CTCCACACCCGTTTAAAGCTCACTATACGAAATCTCTCGTCCGAAACGACGCGGATGTC 420
Db 10234 CTCCACACCCGTTTAAAGCTCACTATACGAAATCTCTCGTCCGAAACGACGCGGATGTC 10293
QY 421 TAGCTAATTTGTCGGTACAGCACCAAGAACCCATTCAGTGTCTTCGGATCCAACTA 480
Db 10294 TAGCTAATTTGTCGGTACAGCACCAAGAACCCATTCAGTGTCTTCGGATCCAACTA 10353
QY 481 TCGGTGTATCAATTCGGGAACACCGCGGACCTCGCGGACTCTATTCAGAGCTTCGTGT 540
Db 10354 TCGGTGTATCAATTCGGGAACACCGCGGACTCTATTCAGAGCTTCGTGT 10413
QY 541 CGCACCTTCGGATTACTACCGTCCAACTTGGGCTCTATTCAGGACCGAGGAAAGTTGG 600
Db 10414 CGCACCTTCGGATTACTACCGTCCAACTTGGGCTCTATTCAGGACCGAGGAAAGTTGG 10473
QY 601 CGCAACTGGCAAGCGTACGTTGCGACGAGGCGCACGACGACCGCGCGGACCAAC 660
Db 10474 CGCAACTGGCAAGCGTACGTTGCGACGAGGCGCACGACGACCGCGCGGACCAAC 10533
QY 661 CCGACGCGCGTCACTGCAACACGCGCTTCGAACTTGAACGCGGAACACTTACCTTCCC 720

Db 10534 CCGAGCCCGCTCACTGCAACACCGAGCGCTTCGAACTTGAACGCGGAACACTTTACCTTTCC 10593
QY 721 TGGCTAGAAAATGGCGTGGATCATTTACGAACCGACACCGCGGAACGAAAATTCAAACGTT 780
Db 10594 TGGCTAGAAAATGGCGTGGATCATTTACGAACCGACACCGCGGAACGAAAATTCAAACGTT 10653
QY 781 ACTGTCGCTCTCGGAGCAATGAGCCCTACGCTAAATTGGGTTAACCGTGGCTGCCGTCGTG 840
Db 10654 ACTGTCGCTCTCGGAGCAATGAGCCCTACGCTAAATTGGGTTAACCGTGGCTGCCGTCGTG 10713
QY 841 AGCGCAACGATCGGCCCTCGCTCAATTGTAATTTCCATCGTCACGAGAAACATGTGCAACCCG 900
Db 10714 AGCGCAACGATCGGCCCTCGCTCAATTGTAATTTCCATCGTCACGAGAAACATGTGCAACCCG 10773
QY 901 CACCGAAAATTAGACACGCTCTCGCAAGACGACGAAACGTTCCCAAACTAGAGGGA 960
Db 10774 CACCGAAAATTAGACACGCTCTCGCAAGACGACGAAACGTTCCCAAACTAGAGGGA 10833
QY 961 TCGGAAAATTTGGACCCATGTTGCTGCGGAAATAAACAAGGCGCTGACCAAGATAGT 1020
Db 10834 TCGGAAAATTTGGACCCATGTTGCTGCGGAAATAAACAAGGCGCTGACCAAGATAGT 10893
QY 1021 GAATTTGTGGAATGTTGCGATTTGTAACCGCTCTGCGCTAAGCTTCGCCCGACTCAATA 1080
Db 10894 GAATTTGTGGAATGTTGCGATTTGTAACCGCTCTGCGCTAAGCTTCGCCCGACTCAATA 10953
QY 1081 AAAATG 1086
Db 10954 AAAATG 10959
RESULT 3
ABK90557
ID ABK90557 standard; DNA; 3605 BP.
XX AC ABK90557;
XX DT 29-AUG-2003 (revised)
XX DT 15-NOV-2002 (first entry)
XX Infectious laryngotracheitis virus (ILTV) DNA.
DE Infectious laryngotracheitis virus; novel recombinant avian herpesvirus;
KW NAHV; herpes virus of turkeys; HVT; Marek's disease; Newcastle disease;
KW infectious laryngotracheitis; virucide; immunostimulant; vaccine; ILTV;
KW gene; ds.
XX Gallid herpesvirus 1.
OS
XX Location/Qualifiers
FH 585..1889
FT /*tag= a
FT /product= "ILTV glycoprotein D"
FT 1997..3085
FT /*tag= b
FT /product= "ILTV glycoprotein I"
XX US2002081316-A1.
XX
XX 27-JUN-2002.
XX
XX 14-JUN-2001; 2001US-00881457.
XX
XX 12-JUN-1992; 92US-00898087.
XX 26-FEB-1993; 93US-00023610.
XX 14-JUN-1993; 93WO-US005681.
XX 09-AUG-1994; 94US-00288065.
XX 09-AUG-1995; 95WO-US010245.
XX 13-JUN-1996; 96US-00863566.
XX 21-FEB-1997; 97US-00804372.
XX 25-OCT-1999; 99US-00426352.
XX (COCH/) COCHRAN M D.


```
FT FT      /tag= f
FT FT      /label= ORF1(RC)
FT FT      /note= "unique-like short 2 (US2) gene"
FT FT      3642..3645
FT FT      /tag= g
FT FT      /note= "TATA signal for ORF1 and ORF2"
FT FT      3675..3678
FT FT      /tag= h
FT FT      /note= "TATA signal for ORF1"
FT FT      3677..3680
FT FT      /tag= i
FT FT      /note= "TATA signal for ORF2"
FT FT      3694..5124
FT FT      /tag= j
FT FT      /label= ORF2
FT FT      /note= "protein kinase gene"
FT FT      5125..5131
FT FT      /tag= k
FT FT      /note= "polyA signal for ORF2"
FT FT      5173..5178
FT FT      /tag= l
FT FT      /note= "polyA signal for ORF2"
FT FT      5210..7081
FT FT      /tag= m
FT FT      /label= ORF3
FT FT      /note= "unique long 47 (UL47)-like gene"
FT FT      6748..7080
FT FT      /tag= o
FT FT      /label= ORF4
FT FT      /note= "complement(6774..7154)"
FT FT      /tag= p
FT FT      /label= ORF4(RC)
FT FT      7158..7161
FT FT      /tag= n
FT FT      /note= "TATA signal for ORF4"
FT FT      7158..7161
FT FT      /tag= q
FT FT      /note= "TATA signal for ORF5"
FT FT      7245..8123
FT FT      /tag= r
FT FT      /label= ORF5
FT FT      /note= "glycoprotein gG gene"
FT FT      7245..7322
FT FT      /tag= s
FT FT      7323..8120
FT FT      /tag= t
FT FT      8200..8205
FT FT      /tag= u
FT FT      /note= "polyA site for ORF5"
FT FT      8289..8294
FT FT      /tag= v
FT FT      /note= "polyA signal for ORF5"
FT FT      8333..11290
FT FT      /tag= w
FT FT      /label= ORF6
FT FT      /note= "glycoprotein g60 gene"
FT FT      complement(9584..10462)
FT FT      /tag= y
FT FT      /label= ORF6(RC)
FT FT      9623..10363
FT FT      /tag= x
FT FT      /note= "repeat region consists of approx. 23 repeats of
FT FT      30-36 bp"
FT FT      11091..11094
FT FT      /tag= z
FT FT      /note= "potential TATA signal for ORF7"
FT FT      11096..12402
FT FT      /tag= aa
FT FT      /label= ORF7
FT FT      12455..12448
FT FT      /tag= ab
FT FT      /note= "TATA signal for ORF8"
FT FT      12510..13598
```

```
FT FT      /tag= ac
FT FT      /label= ORF8
FT FT      /note= "glycoprotein gI gene"
FT FT      12510..12575
FT FT      /tag= ad
FT FT      12576..13595
FT FT      /tag= ae
FT FT      complement(13253..13786)
FT FT      /tag= af
FT FT      /label= ORF8(RC)
FT FT      13705..13708
FT FT      /tag= ag
FT FT      /note= "TATA signal for ORF9"
FT FT      13792..15291
FT FT      /tag= ah
FT FT      /label= ORF9
FT FT      /note= "glycoprotein gE gene"
FT FT      13792..13845
FT FT      /tag= ai
FT FT      13846..15288
FT FT      /tag= aj
FT FT      15116..15119
FT FT      /tag= ak
FT FT      /note= "TATA signal for ORF10"
FT FT      15298..16080
FT FT      /tag= al
FT FT      /label= ORF10
FT FT      16004..18912
FT FT      /tag= am
FT FT      16129..17013
FT FT      /tag= an
FT FT      complement(17380..18216)
FT FT      /tag= ao
FT FT      96WO-US003916.
FT FT      26-SEP-1996.
FT FT      21-MAR-1996; 96WO-US003916.
FT FT      23-MAR-1995; 95US-00410121.
FT FT      06-JUN-1995; 95US-00468190.
FT FT      (SYTR ) SYNTRO CORP.
FT FT      Wild MA, Cochran MD;
FT FT      WPI; 1996-443172/44.
FT FT      P-PSDB; AAW06782, AAW06783, AAW06784, AAW06785, AAW06786, AAW06787,
FT FT      AAW06788, AAW06789, AAW06790, AAW06791, AAW06792, AAW06793, AAW01415,
FT FT      AAW01416.
FT FT      Recombinant infectious laryngotracheitis virus with deletion in the
FT FT      glyco:protein G, gI or US2 gene, etc. - useful for vaccines against
FT FT      infectious laryngotracheitis in poultry.
FT FT      Example 11; Page 138-154; 216pp; English.
FT FT      The nucleotide sequence of 19,912 bp of contiguous DNA (AAU44385) from
FT FT      the unique short and flanking region of infectious laryngotracheitis
FT FT      virus (ILT) contains the entire 13,098 bp unique short region (see also
FT FT      AAT44384) and includes 17 open reading frames that encode proteins
FT FT      (AAW06782-94, AAW01415-16) of over 100 amino acids, 10 of which show
FT FT      homology to other virus genes. Novel recombinant, attenuated ILTV
FT FT      comprises the ILTV genome contg. a deletion in the unique short region,
FT FT      esp. in the glycoprotein gG, gI, US2, ORF4, UL47-like or g60 gene. The
FT FT      attenuated virus is useful as a vaccine against ILTV. A foreign gene
FT FT      encoding an antigen e.g. from another avian virus can be inserted into
FT FT      the US2, UL47-like, ORF4, gG, g60 or gI gene to provide a multivalent
FT FT      vaccine for chickens and other poultry. Deletion of the gG or gI gene
FT FT      provides as a negative marker to distinguish vaccinated from infected
FT FT      animals. (Updated on 16-OCT-2003 to standardise OS field)
```

SQ Sequence 18912 BP; 4416 A; 5261 C; 5251 G; 3984 T; 0 U; 0 Other;
Query Match 99.9%; Score 1084.4; DB 2; Length 18912;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCATCGCTACTTGGAACTCTGGCTCTCTTCCGCGGACGCTCGACCTTCGGGCG 60
DB 12510 ATGGCATCGCTACTTGGAACTCTGGCTCTCTTCCGCGGACGCTCGACCTTCGGGCG 12569

QY 61 ATGGGAATCGTGATCACTGGAAATCACGCTTCCGCGGAGATGACGAGATCAATCGTG 120
DB 12570 ATGGGAATCGTGATCACTGGAAATCACGCTTCCGCGGAGATGACGAGATCAATCGTG 12629

QY 121 ATCGTCGGCGCTCGCGCGGACGCTCAATCAACTGCAGCTATTTTCATCGCTGGCCAG 180
DB 12630 ATCGTCGGCGCTCGCGCGGACGCTCAATCAACTGCAGCTATTTTCATCGCTGGCCAG 12689

QY 181 AGACCCCAAAACCTTACTCAGGAACCGTCGCGGCTTCGCTCTGATATACAAAC 240
DB 12690 AGACCCCAAAACCTTACTCAGGAACCGTCGCGGCTTCGCTCTGATATACAAAC 12749

QY 241 CAGTGTCTACAGGAATTTAGCGAGGCGCTTTGAAAATTCGACTCATCGATCGTCTCT 300
DB 12750 CAGTGTCTACAGGAATTTAGCGAGGCGCTTTGAAAATTCGACTCATCGATCGTCTCT 12809

QY 301 GTTTTGTGCGTGTAAAGTGAACGAGTACAGTTCTCGCGCTCGAAGACGTAACCGGA 360
DB 12810 GTTTTGTGCGTGTAAAGTGAACGAGTACAGTTCTCGCGCTCGAAGACGTAACCGGA 12869

QY 361 CTTCCACACCGCTTAAAGTCACTATAGAAATCTCGTCCGMAAGCAGCGGATGTTTC 420
DB 12870 CTTCCACACCGCTTAAAGTCACTATAGAAATCTCGTCCGMAAGCAGCGGATGTTTC 12929

QY 421 TACGTAATTTGTCGGCTAGACACACCAAGAACCCATTGAGCTCTTCGCGATCCAACTA 480
DB 12930 TACGTAATTTGTCGGCTAGACACACCAAGAACCCATTGAGCTCTTCGCGATCCAACTA 12989

QY 481 TCGGTGTATCAATTCGGAACACCGCGGACTTCGCGGACTCTATTCGAGGCTTCGTTGT 540
DB 12990 TCGGTGTATCAATTCGGAACACCGCGGACTTCGCGGACTCTATTCGAGGCTTCGTTGT 13049

QY 541 CGACCTTCGATTAACCTACGCTCAACTGAGGCTATCTCAGACCGAGGAAGTTGG 600
DB 13050 CGACCTTCGATTAACCTACGCTCAACTGAGGCTATCTCAGACCGAGGAAGTTGG 13109

QY 601 CGCACTGCAAGCTGACTTCCGCGAGGCGGACGACGACGCGCGGCGGACACACC 660
DB 13110 CGCACTGCAAGCTGACTTCCGCGAGGCGGACGACGACGCGCGGCGGACACACC 13169

QY 661 CGAGCGCGCTCACTGCAACGAGCGCTCCGAACTTGAAGCGGAACATTTACCTTTCCC 720
DB 13170 CGAGCGCGCTCACTGCAACGAGCGCTCCGAACTTGAAGCGGAACATTTACCTTTCCC 13229

QY 721 TGGCTAGAAAATGGGCTGATCATTCAGAACCGACACCGGCAACGAAAATTCAAACGTT 780
DB 13230 TGGCTAGAAAATGGGCTGATCATTCAGAACCGACACCGGCAACGAAAATTCAAACGTT 13289

QY 781 ATGTCCGCTCGGACAAATGAGCCCTAGCTTAATTTGGGTHAACCGTGGCTGCGCTG 840
DB 13290 ATGTCCGCTCGGACAAATGAGCCCTAGCTTAATTTGGGTHAACCGTGGCTGCGCTG 13349

QY 841 AGCGCAACGATGGCTCGTCACTGTAATTTCCATCTGTCACGAAACATGTGACCCCG 900
DB 13350 AGCGCAACGATGGCTCGTCACTGTAATTTCCATCTGTCACGAAACATGTGACCCCG 13409

QY 901 CACCGAAAATTAGACGCTCTCGCAAGACGACGAAACGTTTCCCAAACTAGAGGGAA 960
DB 13410 CACCGAAAATTAGACGCTCTCGCAAGACGACGAAACGTTTCCCAAACTAGAGGGAA 13469

QY 961 TCGCGAAAATTTGGACCCCATGTTGCGTGCAGAAATAAACAAGGGGCGCTGACGAGATAGT 1020
DB 13470 TCGCGAAAATTTGGACCCCATGTTGCGTGCAGAAATAAACAAGGGGCGCTGACGAGATAGT 13529

QY 1021 GAACTTGTGAACTGGTTCGATTTGAACCGCTTCGCGTAAAGCTCGCCGACTCAATA 1080
DB 13530 GAACTTGTGAACTGGTTCGATTTGAACCGCTTCGCGTAAAGCTCGCCGACTCAATA 13589

QY 1081 AAAATG 1086
DB 13590 AAAATG 13595

RESULT 5

AAT33513/c
ID AAT33513 standard; DNA; 534 BP.

XX AAT33513;
XX AC
XX 16-OCT-2003 (revised)
DT 19-NOV-1996 (first entry)
XX Infectious laryngotracheitis virus ORF8 reverse complement.
XX Infectious laryngotracheitis virus; ILTV; herpesvirus; attenuation;
KW vector; vaccine; chicken; poultry; immunisation; ds.
XX Gallid herpesvirus 1.

XX W09508622-A1.
XX 30-MAR-1995.
XX 16-SEP-1994; 94WO-US010628.
XX 24-SEP-1993; 93US-00126597.
XX (SVTR) SYNTRO CORP.
XX Wild MA, Cochran MD;
XX WPI; 1995-139591/18.
XX P-PSDB; AAW00642.

Recombinant attenuated infectious laryngotracheitis virus - for use in
vaccines to protect poultry from infection from the virus, also methods
of distinguishing between vaccinated and naturally infected birds.

Example 1; Page 115-116; 177pp; English.

XX ORF8 reverse complement (AAT33513) consists of bases 11,150 to 10,617 on
CC the reverse complement strand of the unique short region (see also
CC AAT33504) of infectious laryngotracheitis virus (ILTV) genomic DNA. It
CC codes for a protein (AAW00642) of approx. 19,470 mol.wt. of unknown
CC function. (Updated on 16-OCT-2003 to standardise OS field)

XX SQ Sequence 534 BP; 115 A; 138 C; 133 G; 148 T; 0 U; 0 Other;

Query Match 31.6%; Score 343; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 5e-99;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 TTACGAACCGACACCGCGCAACGAAATTCAAACGTTACTGTCCGCTCGGGCAATGAG 803
DB 534 TTACGAACCGACACCGCGCAACGAAATTCAAACGTTACTGTCCGCTCGGGCAATGAG 475

QY 804 CCTACGTAATTTGGGTAACCGTGGCTGCGTGGCGCAACGATCGGCTCGTCAAT 863
DB 474 CCTACGCTAATTTGGGTAACCGTGGCTGCGTGGCGCAACGATCGGCTCGTCAAT 415

QY 864 TGTAAATTTCCATCGTCACCAAGAACATGTGACCCCGCACCGAAAATTAGACACGFTTC 923
DB 414 TGTAAATTTCCATCGTCACCAAGAACATGTGACCCCGCACCGAAAATTAGACACGFTTC 355

QY 924 GCAAGACGCAAGAACGTTTCCAAATAGAGGGAATCGGAAAATTTGGACCCATGTT 983

Db	354	GCAAGACGACGAGAACGTTCCAAACTAGAGGGAATCGCGAAATTTGGACCATGGT	295
Qy	984	TGCGTGCAGAAATAAACAAGGGCGCTGACAGGATAGTGAACCTGTGTGAACTGGTTGCGAT	1043
Db	294	TGCGTGCAGAAATAAACAAGGGCGCTGACAGGATAGTGAACCTGTGTGAACTGGTTGCGAT	235
Qy	1044	TGTTAACCCGCTGCGCTGAAGCTCGCCGACTCAATAAAATG	1086
Db	234	TGTTAACCCGCTGCGCTGAAGCTCGCCGACTCAATAAAATG	192
RESULT 6			
ADA71938/c			
ID	ADA71938 standard; DNA; 2000 BP.		
XX			
AC	ADA71938;		
XX			
DT	20-NOV-2003 (first entry)		
XX			
DE	Rice gene, SEQ ID 5263.		
XX			
KW	Plant; bacterial infection; fungal infection; viral infection; rice;		
XX			
OS	Oryza sativa.		
XX			
FN	WO2003000898-A1.		
XX			
PD	03-JAN-2003.		
XX			
PF	22-JUN-2001; 2001WO-IB001105.		
XX			
PR	22-JUN-2001; 2001WO-IB001105.		
XX			
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.		
XX			
PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;		
PI	Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;		
XX			
DR	WPI; 2003-175290/17.		
XX			
PT	Identifying at least one gene involved in plant resistance or response to		
PT	pathogenic infection for conferring resistance or tolerance to a plant to		
PT	bacterial, fungal or viral infection by determining or detecting plant		
PT	gene expression.		
XX			
PS	Claim 27; SEQ ID NO 5263; 899pp; English.		
XX			
CC	The present invention relates to a method (M1) for identifying genes		
CC	involved in plant resistance or response to pathogenic infection. M1		
CC	comprises identifying a gene whose expression is significantly altered in		
CC	the incompatible interaction of plant gene expression relative to		
CC	expression of the gene in an uninfected plant, in a mutant plant that		
CC	does not express a gene associated with response to pathogenic infection,		
CC	or in a corresponding incompatible or compatible interaction. (M1) is		
CC	useful for conferring resistance to resistance or tolerance to a plant to		
CC	bacterial, fungal or viral infection. The present sequence was used to		
CC	illustrate the invention.		
XX			
SQ	Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;		
Query Match 4.4%; Score 47.8; DB 7; Length 2000;			
Best Local Similarity 7.5%; Pred. No. 0.00035;			
Matches 40; Conservative 254; Mismatches 241; Indels 0; Gaps 0;			
Qy	390	AAATCTCTGCGACGACGAGGGGATGTTCTAGTAATTTGCGCTAGACACCAAA	449
Db	854	WWWWRYTMYTCYAMTCAKCKYKAMTKWTTWACAWRATSRWRAWAGRWKRYKMDRA	795
Qy	450	AGAACCCATTGACGCTCTCGGATCCCACTATCGGTGTATCAATTGCGACACCGCCG	509
Db	794	YWWRWRCWAGWAGWAKWRYKWKYATRYTWKMWMTWWSWRKWSYRWWSGWRM	735

Qy	510	GACTCGCGACTCTATTCCAAAGCTTCGTGTCGACACTTCCTCGGATTAATCCTACCGTCCCAACT	569
Db	734	RMSAWRYCSRMKCAKTKYASSARWTGRARSYRFRRWYWKRGWYRYRWRSCTWRA	675
Qy	570	TGAGGCGCTATCTCAGAGACCGAGGAAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGGA	629
Db	674	RMSKREKWAAGASMKSCWYWRGABSMYYSKYSCSAKCKCKTRYMTSSYMTSGMYGMYSSY	615
Qy	630	GGCCACGACGACGCGCGGAGGCGACAACCCCGACGCCCTCACTGCAACCCAGCGCTC	689
Db	614	KSMWSYKMSYMGKMTCTMTYSMGSTRSRKSGRSMGRSMYRWKWKMRKRYNRYWKW	555
Qy	690	CGAECTTGAAGCGGAACACACTTTACCTTTCCCTGGCTAGAAAATGGCGTGGATCATACGA	749
Db	554	KCTWRRCMYRWGYTMYTTSRSRMITGRYKARYTSRRYKRYKRYKRYKRYKRYKRYKRYK	495
Qy	750	ACCGACACCCGCAAACTTCAACGTTTACTGTCGCTCTCGGGAACAATGAGCCCTAC	809
Db	494	SYMRYGYCKACKCCYAMCWAAYSGMMYWKYKSKWMSMSTKYMMSMYKCRSMKY	435
Qy	810	GCTAATGCGGTAAACCGTGGCTGCGCTGAGCGCAACGATCGGCTCGTCATGTGAAT	869
Db	434	GAKGCGCCKWWTYCSYGMKWTYMGYSYKRCYKYMENYMYKGMNMYMYYSAYSSMMTW	375
Qy	870	TTCCATCGTCACCAAGAACATGTCACCCCGCACGAAAATTAGACACGCTCTCG	924
Db	374	YYYAYKYWKYKRRGTMSWYKSYKKYCTWCKYMKCMCYRWKMRKKTYS	320
RESULT 7			
ADA71938			
ID	ADA71938 standard; DNA; 2000 BP.		
XX			
AC	ADA71938;		
XX			
DT	20-NOV-2003 (first entry)		
XX			
DE	Rice gene, SEQ ID 5263.		
XX			
KW	Plant; bacterial infection; fungal infection; viral infection; rice;		
XX			
OS	Oryza sativa.		
XX			
FN	WO2003000898-A1.		
XX			
PD	03-JAN-2003.		
XX			
PF	22-JUN-2001; 2001WO-IB001105.		
XX			
PR	22-JUN-2001; 2001WO-IB001105.		
XX			
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.		
XX			
PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;		
PI	Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;		
XX			
DR	WPI; 2003-175290/17.		
XX			
PT	Identifying at least one gene involved in plant resistance or response to		
PT	pathogenic infection for conferring resistance or tolerance to a plant to		
PT	bacterial, fungal or viral infection by determining or detecting plant		
PT	gene expression.		
XX			
PS	Claim 27; SEQ ID NO 5263; 899pp; English.		
XX			
CC	The present invention relates to a method (M1) for identifying genes		
CC	involved in plant resistance or response to pathogenic infection. M1		
CC	comprises identifying a gene whose expression is significantly altered in		
CC	the incompatible interaction of plant gene expression relative to		
CC	expression of the gene in an uninfected plant, in a mutant plant that		
CC	does not express a gene associated with response to pathogenic infection,		
CC	or in a corresponding incompatible or compatible interaction. (M1) is		
CC	useful for conferring resistance to resistance or tolerance to a plant to		
CC	bacterial, fungal or viral infection. The present sequence was used to		
CC	illustrate the invention.		
XX			
SQ	Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;		
Query Match 4.4%; Score 47.8; DB 7; Length 2000;			
Best Local Similarity 7.5%; Pred. No. 0.00035;			
Matches 40; Conservative 254; Mismatches 241; Indels 0; Gaps 0;			
Qy	390	AAATCTCTGCGACGACGAGGGGATGTTCTAGTAATTTGCGCTAGACACCAAA	449
Db	854	WWWWRYTMYTCYAMTCAKCKYKAMTKWTTWACAWRATSRWRAWAGRWKRYKMDRA	795
Qy	450	AGAACCCATTGACGCTCTCGGATCCCACTATCGGTGTATCAATTGCGACACCGCCG	509
Db	794	YWWRWRCWAGWAGWAKWRYKWKYATRYTWKMWMTWWSWRKWSYRWWSGWRM	735

Qy	510	GACTCGGGACTCTATTCCAAAGGCTTCGTGCGCACTTCGGATTACCTACCGTCCAACT	569
Db	734	RWSAWRYCSRWKCCKTKYASSARWTKRAKRSYRFRRWYKRWKGTYRYRWSCRMTRA	675
Qy	570	TEAGGCTTATCTCAGGACCGAGGAAAGTTGGCGCAACTGCGCAAGCGTAGCTTGCCACGGA	629
Db	674	RMSKERKWAGASMKSCWMYWRGABSWYKYSKSAKCKKTRVYMTSSYMTSGTMYGMYSSY	615
Qy	630	GGCCACGACGACGAGCGCGGAGGCGACAACCCGACGCGCGTCACTGCAACGAGCGCTC	689
Db	614	KMSWTSXMSYMGWMTCTMYTSMKGSSTRSRKMGSRWSRMYRWKWKMRKRKRYMYRWK	555
Qy	690	CGAACTTGAAGCGGAACACTTTACTTTCCTTGCTAGAAAAATGGCGTGATCATATTACGA	749
Db	554	KCTWRRCWYRWGTYTMYTTSRSMYTGKRYKARYTSKRYMYKRYKYCWYTYGMYMKC	495
Qy	750	ACCGACACCCGCAACGAAATTCAAAACGTTACTGTCGCTCGGACCAATGAGCCCTAC	809
Db	494	SYMREYGYCKACKCCYAMCWKAAYSGMMYWKYSKWMRMSTKYMWSWYKCRSMKY	435
Qy	810	GCTAATTGGGTAAACGGTGGCTGCGCTGCTGAGCGCAACGATCGGCTCGCTCATTTAAT	869
Db	434	GAGCYGCKWMTYCSYGYMKWYTMGSYKYSRCYKRYMYMYKGMWYMYYSAYSSMMTW	375
Qy	870	TTCATCTGCACGAGAACATGTGCACCCGCGCACCGGAAATTTAGACACGGTCTCG	924
Db	374	YVYVAKYWKYKRRGTMSWYKSKYKCTWNCYMKMRCYRWKRWKRYKTKYS	320
RESULT 7			
ADA71938			
ID	ADA71938 standard; DNA; 2000 BP.		
XX			
AC	ADA71938;		
XX			
DT	20-NOV-2003 (first entry)		
XX			
DE	Rice gene, SEQ ID 5263.		
XX			
KW	Plant; bacterial infection; fungal infection; viral infection; rice;		
XX			
OS	Oryza sativa.		
XX			
PN	WO2003000898-A1.		
XX			
PD	03-JAN-2003.		
XX			
PF	22-JUN-2001; 2001WO-IB001105.		
XX			
PR	22-JUN-2001; 2001WO-IB001105.		
XX			
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.		
XX			
PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;		
PI	Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;		
XX			
DR	WPI; 2003-175290/17.		
XX			
PT	Identifying at least one gene involved in plant resistance or response to		
PT	pathogenic infection for conferring resistance or tolerance to a plant to		
PT	bacterial, fungal or viral infection by determining or detecting plant		
PT	gene expression.		
XX			
PS	Claim 27; SEQ ID NO 5263; 899pp; English.		
XX			
CC	The present invention relates to a method (M1) for identifying genes		
CC	involved in plant resistance or response to pathogenic infection. M1		
CC	comprises identifying a gene whose expression is significantly altered in		
CC	the incompatible interaction of plant gene expression relative to		
CC	expression of the gene in an uninfected plant, in a mutant plant that		
CC	does not express a gene associated with response to pathogenic infection,		
CC	or in a corresponding incompatible or compatible interaction. (M1) is		

XX Sequence 780 BP; 73 A; 226 C; 314 G; 167 T; 0 U; 0 Other;
SQ Query Match 3.5%; Score 38.2; DB 7; Length 780;
Best Local Similarity 58.3%; Pred. No. 0.26; 48; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 29 TCCTTGGCGGACGCTCGACCCCTTCGGGGCGATGGGAATCGTGATCACTGGAATCAGC 88
DB 544 TCCACGCCGCGACGCTGGGACCCCTCGGTGGCGGACGATCACCCTCGGTGACGG 485
QY 89 TCTCCGCCGAGGATTTGACGACGATCAATCGTATCGTCGGCTCGCCCGGAGC 143
DB 484 TCGGCCCCGAGGACGCTGTCCAGCGGCGGAGGACACCGCGGCGGCGGC 430

RESULT 10
AAI99682_06/c
Continuation (7 of 45) of AAI99682 from base 600001 (Mycobacterium tuberculosis strain 1
WP Sequence split into 45 fragments LOCUS AAI99682 Accession Aai99682
WP Fragment Name Begin End
WP AAI99682_00 1 110000
WP AAI99682_01 100001 210000
WP AAI99682_02 200001 310000
WP AAI99682_03 300001 410000
WP AAI99682_04 400001 510000
WP AAI99682_05 500001 610000
WP AAI99682_06 600001 710000
WP AAI99682_07 700001 810000
WP AAI99682_08 800001 910000
WP AAI99682_09 900001 1010000
WP AAI99682_10 1000001 1110000
WP AAI99682_11 1100001 1210000
WP AAI99682_12 1200001 1310000
WP AAI99682_13 1300001 1410000
WP AAI99682_14 1400001 1510000
WP AAI99682_15 1500001 1610000
WP AAI99682_16 1600001 1710000
WP AAI99682_17 1700001 1810000
WP AAI99682_18 1800001 1910000
WP AAI99682_19 1900001 2010000
WP AAI99682_20 2000001 2110000
WP AAI99682_21 2100001 2210000
WP AAI99682_22 2200001 2310000
WP AAI99682_23 2300001 2410000
WP AAI99682_24 2400001 2510000
WP AAI99682_25 2500001 2610000
WP AAI99682_26 2600001 2710000
WP AAI99682_27 2700001 2810000
WP AAI99682_28 2800001 2910000
WP AAI99682_29 2900001 3010000
WP AAI99682_30 3000001 3110000
WP AAI99682_31 3100001 3210000
WP AAI99682_32 3200001 3310000
WP AAI99682_33 3300001 3410000
WP AAI99682_34 3400001 3510000
WP AAI99682_35 3500001 3610000
WP AAI99682_36 3600001 3710000
WP AAI99682_37 3700001 3810000
WP AAI99682_38 3800001 3910000
WP AAI99682_39 3900001 4010000
WP AAI99682_40 4000001 4110000
WP AAI99682_41 4100001 4210000
WP AAI99682_42 4200001 4310000
WP AAI99682_43 4300001 4410000
WP AAI99682_44 4400001 4411529

Query Match 3.5%; Score 38.2; DB 4; Length 110000;
Best Local Similarity 58.3%; Pred. No. 2.9;
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 29 TCCTTGGCGGACGCTCGACCCCTTCGGGGCGATGGGAATCGTGATCACTGGAATCAGC 88
DB 544 TCCACGCCGCGACGCTGGGACCCCTCGGTGGCGGACGATCACCCTCGGTGACGG 485
QY 89 TCTCCGCCGAGGATTTGACGACGATCAATCGTATCGTCGGCTCGCCCGGAGC 143
DB 484 TCGGCCCCGAGGACGCTGTCCAGCGGCGGAGGACACCGCGGCGGCGGC 430

RESULT 9
ACA40341/c
ID ACA40341 standard; DNA; 780 BP.
XX ACA40341;
AC ACA40341;
DT 19-JUN-2003 (first entry)
DE Prokaryotic essential gene #21998.
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX Mycobacterium tuberculosis.
OS
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR F-PSDB; ABU36471.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 28211; 1765pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in an
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

Db 18034 TCCACGGCCACGCTGCGCACCTCGGTGGCGGAGCAACGGTGATCACCCCGTCAGCG 17975

Qy 89 TCTCCGCGCAGGATTGACGACGATCACATCGGTGCGCGCTCGCCCGGAGC 143

Db 17974 TCGGCCCGCAGGACGCTGTCTCCAGCGGAGCGGAGAACCGCGCAAGCACCGCGGC 17920

RESULT 11

AAI99683_06/c

Continuation (7 of 44) of AAI99683 from base 600001 (Mycobacterium tuberculosis strain H

WP Sequence split into 44 fragments LOCUS AAI99683 Accession Aai99683

Fragment Name Begin End

WP AAI99683_00 1 110000

WP AAI99683_01 100001 210000

WP AAI99683_02 200001 310000

WP AAI99683_03 300001 410000

WP AAI99683_04 400001 510000

WP AAI99683_05 500001 610000

WP AAI99683_06 600001 710000

WP AAI99683_07 700001 810000

WP AAI99683_08 800001 910000

WP AAI99683_09 900001 1010000

WP AAI99683_10 1000001 1110000

WP AAI99683_11 1100001 1210000

WP AAI99683_12 1200001 1310000

WP AAI99683_13 1300001 1410000

WP AAI99683_14 1400001 1510000

WP AAI99683_15 1500001 1610000

WP AAI99683_16 1600001 1710000

WP AAI99683_17 1700001 1810000

WP AAI99683_18 1800001 1910000

WP AAI99683_19 1900001 2010000

WP AAI99683_20 2000001 2110000

WP AAI99683_21 2100001 2210000

WP AAI99683_22 2200001 2310000

WP AAI99683_23 2300001 2410000

WP AAI99683_24 2400001 2510000

WP AAI99683_25 2500001 2610000

WP AAI99683_26 2600001 2710000

WP AAI99683_27 2700001 2810000

WP AAI99683_28 2800001 2910000

WP AAI99683_29 2900001 3010000

WP AAI99683_30 3000001 3110000

WP AAI99683_31 3100001 3210000

WP AAI99683_32 3200001 3310000

WP AAI99683_33 3300001 3410000

WP AAI99683_34 3400001 3510000

WP AAI99683_35 3500001 3610000

WP AAI99683_36 3600001 3710000

WP AAI99683_37 3700001 3810000

WP AAI99683_38 3800001 3910000

WP AAI99683_39 3900001 4010000

WP AAI99683_40 4000001 4110000

WP AAI99683_41 4100001 4210000

WP AAI99683_42 4200001 4310000

WP AAI99683_43 4300001 4403765

Query Match 3.5%; Score 38.2; DB 4; Length 110000;

Best Local Similarity 58.3%; Pred. No. 2.9; Mismatches 48; Indels 0; Gaps 0;

Matches 67; Conservative 0;

Qy 29 TCTTCCGCGCAGCTGCGCACCTTCGGCGCGATGGGAATCGTGATCGTGCCTCGCCCGGAGC 88

Db 19394 TCCACGCGCCAGCTGGCGACCTCGGTGGCGGAGCAACGGTGATCACCCCGGTCAGCG 19335

Qy 89 TCTCCGCCAGGATTGACGACGATCACATCGTGCCTCGCGCTCGCCCGGAGC 143

Db 19334 TCGGCCCGCAGGACGCTGTCCAGCGGAGCGGAGAACACCGCGCAAGCACCGCGGC 19280

RESULT 12

AAF26455

ID AAF26455 standard; cDNA; 2526 BP.

XX AAF26455;

AC 26-MAR-2001 (first entry)

XX D.melanogaster corkscrew DNA.

DE SHP-2; SHP-1; Src Homology-2; protein tyrosine phosphatase; mutant;

XX neoplastic disorder; obesity; angiogenesis; cancer; immune;

KW hematopoietic; allergy; ss.

XX Drosophila melanogaster.

XX US6156551-A.

PN 05-DEC-2000.

XX 05-JUN-1998; 98US-00092443.

XX 05-JUN-1998; 98US-00092443.

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

PA (JOSL-) JOSLIN DIABETES CENT.

XX Neel BG, Shoelson S, Pluskey S, O'reilly AM;

XX WPI; 2001-060166/07.

XX Mutant SH2 domain-containing protein tyrosine phosphatase, useful in in vitro assays to screen for binding partners, inhibitors of tyrosine phosphatase and for treating tyrosine phosphatase-mediated diseases.

XX Disclosure; Fig 7; 161pp; English.

XX The present invention relates to an activated SH2 (Src Homology-2) - domain containing protein tyrosine phosphatase (SHP-1 or SHP-2) mutant with a mutation in the SH2-domain. Activated mutants of SH2-domain-containing protein tyrosine phosphatases are useful in in vitro assays to screen for binding partners and inhibitors of the phosphatase and in the treatment of PTP-mediated diseases or conditions in a mammal, including neoplastic disorders, obesity and to inhibit angiogenesis. Inhibitors identified using the activated mutants are useful for the treatment of cancer, immunosuppression, immunostimulation, hematopoietic stimulation and anti-allergy treatment

XX SQ Sequence 2526 BP; 616 A; 742 C; 732 G; 436 T; 0 U; 0 Other;

Query Match 3.5%; Score 38; DB 5; Length 2526;

Best Local Similarity 51.8%; Pred. No. 0.53;

Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 569 TTGAGGCTATCTCAGACCGGAGGAAAGTTGGCGCACTGGCAAGCGGTACGTTGCCACGG 628

Db 1325 TTAAGACCTACATCGCCACCCAGGCTGTCTGTCTACCCAGCAAGTGAACACCGTGACGG 1384

Qy 629 AGGCCAGGACCGCGCGGAGCGGAGCAACCCCGACGCCCTCACTGCAACACCGGCGCT 688

Db 1385 ACTTCTGGAACATGGTCTGCGACGAGAACACCGGGGTGATCTGTCATGACCCACAGAGT 1444

Qy 689 CCGAACCTTGAAGCGGAACTTTACCTTTCCCTGGGTAGAAATGG 734

Db 1445 ACAGCGCGCAAGAAAGTGGCGGCTACTTGGCGGACGAGG 1490

RESULT 13

ABL06927

ID ABL06927 standard; cDNA; 2677 BP.

XX ABL06927;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 15263.

```
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB62824.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 15263; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2677 BP; 670 A; 760 C; 749 G; 498 T; 0 U; 0 Other;
XX
Query Match 3.5%; Score 38; DB 4; Length 2677;
Best Local Similarity 51.8%; Pred. No. 0.55;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 569 TTGAGGCCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGG 628
DB 1464 TTAAGACCTACATCGCCACCCAGGGCTGTCTGCTACCCAGCAAGTGAACACGGTGACGG 1523
QY 629 AGGCCAGCAGCAGCGCCGAGCGGACACCCCGAGCCCGCTCACTGCAACCGAGCGCT 688
DB 1524 ACTTCTGGAACATGGTCTGGCAGGAGAAACACGCGGGTGATCGTCATGACCAACGAGGAGT 1583
QY 689 CCGNACTTGAAGCGGACACTTTACCTTTCCCTGGCTAGAAATGG 734
DB 1584 ACGAGCGCGCAAGAAAGTGGCCCGCTACTGCGCGGACGAGGG 1629
RESULT 14
ABL18521
ID ABL18521 standard; DNA; 2838 BP.
XX ABL18521;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 7036.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB62824.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 7036; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2677 BP; 670 A; 760 C; 749 G; 498 T; 0 U; 0 Other;
XX
Query Match 3.5%; Score 38; DB 4; Length 2677;
Best Local Similarity 51.8%; Pred. No. 0.55;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 569 TTGAGGCCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGG 628
DB 1464 TTAAGACCTACATCGCCACCCAGGGCTGTCTGCTACCCAGCAAGTGAACACGGTGACGG 1523
QY 629 AGGCCAGCAGCAGCGCCGAGCGGACACCCCGAGCCCGCTCACTGCAACCGAGCGCT 688
DB 1524 ACTTCTGGAACATGGTCTGGCAGGAGAAACACGCGGGTGATCGTCATGACCAACGAGGAGT 1583
QY 689 CCGNACTTGAAGCGGACACTTTACCTTTCCCTGGCTAGAAATGG 734
DB 1584 ACGAGCGCGCAAGAAAGTGGCCCGCTACTGCGCGGACGAGGG 1629
RESULT 15
ABL21135
ID ABL21135 standard; DNA; 4528 BP.
XX ABL21135;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 14878.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
```

Search completed: March 7, 2004, 23:09:33
Job time : 556 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 22:43:59 ; Search time 120 Seconds
(without alignments)
5022.306 Million cell updates/sec

Title: US-09-993-777-1_COPY_9874_10959

Perfect score: 1086
Sequence: 1 ATGGCATCGTACTTGGAC.....CGCCGACTCAATAAAATG 1086

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1086	100.0	13473	5	PCT-US96-03916-1
2	1084.4	99.9	18912	5	PCT-US96-03916-59
3	343	31.6	534	5	PCT-US96-03916-18
4	38.2	3.5	4403765	3	US-09-103-840A-2
5	38.2	3.5	4411529	3	US-09-103-840A-1
6	35.6	3.3	420	4	US-09-252-991A-4761
7	35.6	3.3	642	4	US-09-252-991A-4683
8	35.6	3.3	960	4	US-09-252-991A-4705
9	35.6	3.3	1554	4	US-09-252-991A-4723
10	35.2	3.2	19718	4	US-08-961-527-99
11	34.6	3.2	50341	1	US-08-247-901C-1
12	34.6	3.2	50341	1	US-09-075-904-1
13	34.6	3.2	52297	3	US-09-426-436-1
14	34.6	3.2	52297	3	US-08-705-557-1
15	33.2	3.1	1230025	4	US-09-198-452A-1
16	33	3.0	1068	4	US-09-489-039A-689
17	32.6	3.0	2010	4	US-09-527-972-3
18	32.4	3.0	3024	4	US-09-833-381-889
19	32.2	3.0	4403765	3	US-09-103-840A-2
20	32.2	3.0	4411529	3	US-09-103-840A-1
21	31.8	2.9	1145	3	US-08-809-103B-7
22	31.8	2.9	1148	3	US-08-809-103B-1
23	31.8	2.9	1150	3	US-08-809-103B-3
24	31.8	2.9	1150	3	US-08-809-103B-5
25	31.8	2.9	7218	1	US-08-232-463-14
26	31.6	2.9	843	3	US-08-513-974B-375
27	31.6	2.9	25165	4	US-09-453-702B-39

C	28	30.8	2.8	948	4	US-09-134-001C-1418	Sequence 1418, Ap
	29	30.8	2.8	1728	1	US-08-427-097-1	Sequence 1, Appl
	30	30.8	2.8	1728	2	US-08-878-957-1	Sequence 1, Appl
	31	30.8	2.8	1746	4	US-09-489-039A-4325	Sequence 4325, Ap
	32	30.8	2.8	1752	1	US-08-427-097-15	Sequence 15, Appl
	33	30.8	2.8	1752	1	US-08-427-097-27	Sequence 27, Appl
	34	30.8	2.8	1752	2	US-08-878-957-15	Sequence 15, Appl
	35	30.8	2.8	1752	2	US-08-878-957-27	Sequence 27, Appl
C	36	30.6	2.8	245	4	US-09-313-294A-3607	Sequence 3607, Ap
	37	30.6	2.8	505	4	US-09-621-976-15639	Sequence 15639, A
C	38	30.6	2.8	795	4	US-09-252-991A-4788	Sequence 4788, Ap
	39	30.6	2.8	1047	4	US-09-252-991A-14678	Sequence 14678, A
	40	30.6	2.8	1512	4	US-09-252-991A-1320	Sequence 1320, Ap
	41	30.6	2.8	1614	4	US-09-616-289-45	Sequence 45, Appl
	42	30.6	2.8	1686	4	US-09-252-991A-14548	Sequence 14548, A
C	43	30.6	2.8	1815	4	US-09-252-991A-15077	Sequence 15077, A
C	44	30.6	2.8	2130	4	US-09-252-991A-1148	Sequence 1148, Ap
	45	30.6	2.8	12425	4	US-09-616-289-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1
PCT-US96-03916-1
; Sequence 1, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13473 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1059..2489
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2575..4107
; FEATURE:

```
NAME/KEY: CDS
LOCATION: 4113..4445
FEATURE:
NAME/KEY: CDS
LOCATION: 4609..5487
FEATURE:
NAME/KEY: CDS
LOCATION: 5697..8654
FEATURE:
NAME/KEY: CDS
LOCATION: 9874..10962
FEATURE:
NAME/KEY: CDS
LOCATION: 11159..12658
FEATURE:
NAME/KEY: CDS
LOCATION: 12665..13447
PCT-US96-03916-1

Query Match      100.0%; Score 1086; DB 5; Length 13473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1086; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCATCGCTACTTTGGAACCTGCGCTCTCTCTGCGCGAGCGTCGACCCCTTCGGGCGG 60
DB 9874 ATGGCATCGCTACTTTGGAACCTGCGCTCTCTCTGCGCGAGCGTCGACCCCTTCGGGCGG 9933

QY 61 ATGGGAATCGTGATCACTGGAATCACTCTCCGCGAGGATGAGAGCATCATCGTG 120
DB 9934 ATGGGAATCGTGATCACTGGAATCACTCTCCGCGAGGATGAGAGCATCATCGTG 9993

QY 121 ATCGTCGCGCTCGCGCGAGCTACATTAATCACTGCGAGCTATTTTCATGCGCTGGCGAG 180
DB 9994 ATCGTCGCGCTCGCGCGAGCTACATTAATCACTGCGAGCTATTTTCATGCGCTGGCGAG 10053

QY 181 AGACCCCAACAACTCTACTGAGAACCTGCGCGTTCGCTGCTGATATACAAAC 240
DB 10054 AGACCCCAACAACTCTACTGAGAACCTGCGCGTTCGCTGCTGATATACAAAC 10113

QY 241 CAGTGCTACCGAGAACTTAGCGAGGAGCGCTTTGAAATTCACATCGATCGCTTCT 300
DB 10114 CAGTGCTACCGAGAACTTAGCGAGGAGCGCTTTGAAATTCACATCGATCGCTTCT 10173

QY 301 GTTTTGTGCGGTGTAAGTAGTACCGAGTACAGTTCTCGGCTCGAACAGACTAACCGGA 360
DB 10174 GTTTTGTGCGGTGTAAGTAGTACCGAGTACAGTTCTCGGCTCGAACAGACTAACCGGA 10233

QY 361 CTTCCACACCGCTTTAAGCTCACTATACGAATCTGTCGAGAGCGAGCGAGATGTC 420
DB 10234 CTTCCACACCGCTTTAAGCTCACTATACGAATCTGTCGAGAGCGAGCGAGATGTC 10293

QY 421 TACGTAATTTGTCGCTAGACACACCAAGAACCACTTGACGCTTCGCGATCCAACTA 480
DB 10294 TACGTAATTTGTCGCTAGACACACCAAGAACCACTTGACGCTTCGCGATCCAACTA 10353

QY 481 TCGGTGTATCAATTCGGAACACCGCGAGCTCGCGAGCTCTATTCGAAGCTTCGTGT 540
DB 10354 TCGGTGTATCAATTCGGAACACCGCGAGCTCGCGAGCTCTATTCGAAGCTTCGTGT 10413

QY 541 CGCACCTTCGGATTAACCTACCGTCAACTTGAAGGCTATCTCAGACCGAGGAAGTTGG 600
DB 10414 CGCACCTTCGGATTAACCTACCGTCAACTTGAAGGCTATCTCAGACCGAGGAAGTTGG 10473

QY 601 CGCACTGCGAGCGTGTGTCGAGGAGCGGCGAGGACGAGCGCGGAGCGAGCAAC 660
DB 10474 CGCACTGCGAGCGTGTGTCGAGGAGCGGCGAGGACGAGCGCGGAGCGAGCAAC 10533

QY 661 CGACGCGCGTCACTGCAACGAGCGCTCCGAATTTGAAGCGGAACATTTACCTTTCCC 720
DB 10534 CGACGCGCGTCACTGCAACGAGCGCTCCGAATTTGAAGCGGAACATTTACCTTTCCC 10593

QY 721 TGGCTAGAAAATGGCGTGGATCATTTACGAACCGACCGCGCAACGAAAATTCAAACGTT 780
```

```
DB 10594 TGGCTAGAAAATGGCGTGGATCATTTACGAACCGACACCGCGAAACGAAAATTCAAACGTT 10653
QY 781 ACTGTCGCGTCTCGGACCAATGAGCCCTACGCTAAATTGGGTAACCGTGGCTGCCGTG 840
DB 10654 ACTGTCGCGTCTCGGACCAATGAGCCCTACGCTAAATTGGGTAACCGTGGCTGCCGTG 10713
QY 841 AGCGCAACGATCGGCTCGTCAATTGTAATTTCCATCGTCAACGAAACATGTGCAACCCG 900
DB 10714 AGCGCAACGATCGGCTCGTCAATTGTAATTTCCATCGTCAACGAAACATGTGCAACCCG 10773
QY 901 CACCGAAAATTTAGACACCGTCTCGCAAGACGAGAACGTTCCCAACTAGAGGGAA 960
DB 10774 CACCGAAAATTTAGACACCGTCTCGCAAGACGAGAACGTTCCCAACTAGAGGGAA 10833
QY 961 TCGCGAAAATTTGGACCCCATGTTGCGTGCAGAAATAAACAAGGCGCTGACCAAGAT 1020
DB 10834 TCGCGAAAATTTGGACCCCATGTTGCGTGCAGAAATAAACAAGGCGCTGACCAAGAT 10893
QY 1021 GAATTTGTGGAACCTGTTGCGATTTGTAACCGTCTGCGCTAAGCTCGCCCGACTCAATA 1080
DB 10894 GAATTTGTGGAACCTGTTGCGATTTGTAACCGTCTGCGCTAAGCTCGCCCGACTCAATA 10953
QY 1081 AAAATG 1086
DB 10954 AAAATG 10959

RESULT 2
PCT-US96-03916-59
Sequence 59, Application PC/TUS9603916
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 18912 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: CDS
```


STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/136,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..533
PCT-US96-03916-18

Query Match 31.6%; Score 343; DB 5; Length 534;
Best Local Similarity 100.0%; Pred. No. 4.9e-101;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 TTACGACCGACACCGCGCAAAATCAAACTTACTGCGCTCGGACCAATGAG 803
DB 534 TTACGACCGACACCGCGCAAAATCAAACTTACTGCGCTCGGACCAATGAG 475
QY 804 CCCTAGCTAAATGGGGTAAACCGTGGCTGCGTGGAGCAACGATCGGCTCGTCAT 863
DB 474 CCCTAGCTAAATGGGGTAAACCGTGGCTGCGTGGAGCAACGATCGGCTCGTCAT 415
QY 864 TGTAATTCATCGTCACAGAAACATGTGCACCCCGCACGAAATTAGACACGCTCTC 923
DB 414 TGTAATTCATCGTCACAGAAACATGTGCACCCCGCACGAAATTAGACACGCTCTC 355
QY 924 GCAAGACGACGAGAACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCATGTT 983
DB 354 GCAAGACGACGAGAACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCATGTT 295
QY 984 TGGGTGCGAATAAACAAGGCGCTGACACGAGTACTGAACTTGTGGAATGTTGGCGAT 1043
DB 294 TGGGTGCGAATAAACAAGGCGCTGACACGAGTACTGAACTTGTGGAATGTTGGCGAT 235
QY 1044 TGTTAAACCCGCTGCGCTAGCTCGCCCGCACTCAATAAATAATG 1086
DB 234 TGTTAAACCCGCTGCGCTAGCTCGCCCGCACTCAATAAATAATG 192

RESULT 4
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/136,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..533
PCT-US96-03916-18

Query Match 31.6%; Score 343; DB 5; Length 534;
Best Local Similarity 100.0%; Pred. No. 4.9e-101;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 TTACGACCGACACCGCGCAAAATCAAACTTACTGCGCTCGGACCAATGAG 803
DB 534 TTACGACCGACACCGCGCAAAATCAAACTTACTGCGCTCGGACCAATGAG 475
QY 804 CCCTAGCTAAATGGGGTAAACCGTGGCTGCGTGGAGCAACGATCGGCTCGTCAT 863
DB 474 CCCTAGCTAAATGGGGTAAACCGTGGCTGCGTGGAGCAACGATCGGCTCGTCAT 415
QY 864 TGTAATTCATCGTCACAGAAACATGTGCACCCCGCACGAAATTAGACACGCTCTC 923
DB 414 TGTAATTCATCGTCACAGAAACATGTGCACCCCGCACGAAATTAGACACGCTCTC 355
QY 924 GCAAGACGACGAGAACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCATGTT 983
DB 354 GCAAGACGACGAGAACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCATGTT 295
QY 984 TGGGTGCGAATAAACAAGGCGCTGACACGAGTACTGAACTTGTGGAATGTTGGCGAT 1043
DB 294 TGGGTGCGAATAAACAAGGCGCTGACACGAGTACTGAACTTGTGGAATGTTGGCGAT 235
QY 1044 TGTTAAACCCGCTGCGCTAGCTCGCCCGCACTCAATAAATAATG 1086
DB 234 TGTTAAACCCGCTGCGCTAGCTCGCCCGCACTCAATAAATAATG 192

RESULT 4
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 3.5%; Score 38.2; DB 3; Length 4403765;
Best Local Similarity 58.3%; Pred. No. 3.6;
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 29 TCCTTGGCGGACGCTCGCACCCCTTCGGCGGATGGGAATCGTGATCACTGGAATCAG 88
DB 619394 TCCACGCGCCACGCTGCGCACCCCTTCGGTGGCGGACGCGTATCACCCCGTCAGCG 619335
QY 89 TCTCCGCGCAGATTGACGACGATCACAATCGTGATCGTGGCGCTCGCCCGCAAGC 143
DB 619334 TCGGCCCCAGGACGCGTGTCCAGCGGACGCGCAACCCGCGCAAGCAGCGCGC 619280

RESULT 5
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match 3.5%; Score 38.2; DB 3; Length 4411529;
Best Local Similarity 58.3%; Pred. No. 3.6;
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 29 TCCTTGGCGGACGCTCGCACCCCTTCGGCGGATGGGAATCGTGATCACTGGAATCAG 88
DB 618034 TCCACGCGCCACGCTGCGCACCCCTTCGGTGGCGGACGCGTATCACCCCGTCAGCG 617975
QY 89 TCTCCGCGCAGATTGACGACGATCACAATCGTGATCGTGGCGCTCGCCCGCAAGC 143
DB 617974 TCGGCCCCAGGACGCGTGTCCAGCGGACGCGCAACCCGCGCAAGCAGCGCGC 617920

RESULT 6
US-09-252-991A-4761/c
; Sequence 4761, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: MARC J. RUBENFIELD ET AL.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4761
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4761

Query Match 3.3%; Score 35.6; DB 4; Length 420;
Best Local Similarity 51.2%; Pred. No. 0.16;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 508 GCGACTCGGAGCTATTTCAGGCTTCGTGTCGACCTTCGGATTACCTACCGTCCAA 567
DB 276 GCGCGCGCGCGCTGCGCACCTGTTCTCGGCCATCCGCGACCGACCGCAAC 217
QY 568 CTTGAGGCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGGCAAGCTACGTGCCACG 627
DB 216 GCGCGCGCGCTTGACCGCGCGACGAGGCGGCTGCTGGAGGAGTCCGCGCCAGG 157
QY 628 GAGGCCACGACGACGCGCGGCGGACAAACCCCGACGCC 669
DB 156 GTGGCGCCAGCGAGTACTCGAGGAGCGCGACCTCGCGCGC 115

RESULT 7

US-09-252-991A-4683
; Sequence 4683, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4683

; LENGTH: 642

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4683

Query Match 3.3%; Score 35.6; DB 4; Length 642;
Best Local Similarity 51.2%; Pred. No. 0.2;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 508 GCGACTCGGAGCTATTTCAGGCTTCGTGTCGACCTTCGGATTACCTACCGTCCAA 567
DB 280 GCGCGCGCGCGCTGCGCACCTGTTCTCGGCCATCCGCGACCGACCGCAAC 339
QY 568 CTTGAGGCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGGCAAGCTACGTGCCACG 627
DB 340 GCGCGCGCGCTTGACCGCGCGACGAGCGCGCTCTGAGGAGCTCCGCGCCAGG 399
QY 628 GAGGCCACGACGACCGCGGCGGACAAACCCCGACGCC 669
DB 400 GTGGCGCCAGCGAGTACTCGAGGAGCGCGACCTTCGCGCGC 441

RESULT 8

US-09-252-991A-4705
; Sequence 4705, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4705
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4705

Query Match 3.3%; Score 35.6; DB 4; Length 960;
Best Local Similarity 51.2%; Pred. No. 0.25; 79; Indels 0; Gaps 0;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 508 GCGACTCGGAGCTATTTCAGGCTTCGTGTCGACCTTCGGATTACCTACCGTCCAA 567
DB 213 GCGCGCGCGCGCTGCGCACCTGTTCTCGGCCATCCGCGACCGACCGCAAC 272
QY 568 CTTGAGGCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGGCAAGCTACGTGCCACG 627
DB 273 GCGCGCGCGCTTGACCGCGCGACGAGCGCGCTGCTGGAGGAGCTCCGCGCCAGG 332
QY 628 GAGGCCACGACGACGCGCGGCGGACAAACCCCGACGCC 669
DB 333 GTGGCGCCAGCGAGTACTCGAGGAGCGCGACCTTCGCGCGC 374

RESULT 9

US-09-252-991A-4723
; Sequence 4723, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4723

; LENGTH: 1554

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4723

Query Match 3.3%; Score 35.6; DB 4; Length 1554;
Best Local Similarity 51.2%; Pred. No. 0.33;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 508 GCGACTCGGAGCTATTTCAGGCTTCGTGTCGACCTTCGGATTACCTACCGTCCAA 567
DB 1112 GCGCGCGCGCGCTGCGCACCTGTTCTTCGCCCATCCGCGACCGACCGCAAC 1171
QY 568 CTTGAGGCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGGCAAGCTACGTGCCACG 627
DB 1172 GCGCGCGCGCTTGACCGCGCGACGAGCGCGCTGCTGGAGGAGCTCCGCGCCAGG 1231
QY 628 GAGGCCACGACGACCGCGCGGCGGACAAACCCCGACGCC 669

DB 1232 GTGGCGCAGCGAGTACTCTGAGGACGCCGACCTGCGCGGC 1273

STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,901C
FILING DATE: May 23, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/057,531
FILING DATE: April 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/273
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 50341
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: L5 shuttle phasmid sequence
HYPOTHETICAL: NO
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: L5 mycobacteriophage
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION: No. 5750384e
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-247-901C-1

Query Match 3.2%; Score 34.6; DB 1; Length 50341;
Best Local Similarity 49.2%; Pred. No. 4.6;
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0

431 TTGCGCTAGACACCAAGAACCCATTGACGCTTCGCGATCCAACTATCGGTATC 490

1232 GTGGCGCAGCGAGTACTCTGAGGACGCCGACCTGCGCGGC 1273

RESULT 10
US-08-961-527-99
Sequence 99, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 19718 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-99

Query Match 3.2%; Score 35.2; DB 4; Length 19718;
Best Local Similarity 50.0%; Pred. No. 1.8;
Matches 88; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

664 AGCCCGCTCAGTGCACACGCGCTCCGAACTTGAGCGGGAACACTTACCTTCCCTGG 723
3316 ACCCAATAATACACCGCAGAAACACCGATAACCAAGTTAATCACTGTGCGAATACAGA 3375

724 CTAGAAATGGGTGATCATTAACGACCGACACCGCAACGAAATTCAGAGTTACT 783
3376 AATGAGGTGAGGTACGAGCTCCGAAACGACACCGCTCAAGAGCGATTTACCGTTACT 3435

784 GTCCGCTCCGGACATAGCGCTACGCTAATTTGGGGTAACCGTGGCTGCGGTGT 839
3436 GTCACTACCGAACCAATGCTCCGCAATTTGGCTTTGATATAACGACACTAAGTGT 3491

RESULT 11
US-08-247-901C-1/c
Sequence 1, Application US/08247901C
Patent No. 5750384
GENERAL INFORMATION:
APPLICANT: Jacobs et al
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:

Db 18665 TTCGGCTGTTGAGGATGACGACCCCAAGACGATGTCTCGCTTCAGGGCTCGGTGGAC 18606
Qy 491 AATTGCGGACACCGCCGCGACTTCGCGACTCTATTCCAAGGCTTCGTGCGACCTTCG 550
Db 18605 AAGTAGCGAGCGCGGGGTAGTTCGCCGCTCTTCAACAACGACCTTGACGGAGGTG 18546
Qy 551 GATTACCTACGTCGCACTTGAAGGCTTCTCAGGACCGGAGAAAGTTGGCGCACTGGC 610
Db 18545 TAGAACAACCCCTCCACGCTGTGTGCCAGGTACACACCCCTGGTCACCGGTGCTCAAGTTG 18486
Qy 611 AAGCG 615
Db 18485 AAACG 18481

RESULT 12

US-09-075-904-1/c
; Sequence 1, Application US/09075904
; Patent No. 5994137

GENERAL INFORMATION:

APPLICANT: Jacobs, et al.
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/075,904
FILING DATE: May 11, 1998

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,901
FILING DATE: May 23, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/475
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 50341
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: L5 shuttle phasmid sequence
HYPOTHETICAL: No
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: L5 mycobacteriophage
STRAIN:

INDIVIDUAL ISOLATE:

DEVELOPMENTAL STAGE:

HAPLOTYPE:

TISSUE TYPE:

CELL TYPE:

CELL LINE:

ORGANELLE:

IMMEDIATE SOURCE:

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

FEATURE:

NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION: No. 5994137e
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:

US-09-075-904-1

Query Match 3.2%; Score 34.6; DB 2; Length 50341;
Best Local Similarity 49.2%; Pred. No. 4.6;
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Qy 431 TTCGGCTAGACGACACCAAGAACCCATTGACGTCTTCGCGATCCGAACCTATCGGTATC 490
Db 18665 TTCGGCTGTTGAGGATGACGACCCCAAGACGATGTCTCGCTTCAGGGCTCGGTGGAC 18606
Qy 491 AATTGCGGACACCGCCGCGACTTCGCGACTCTATTCCAAGGCTTCGTGCGACCTTCG 550
Db 18605 AAGTAGCGAGCGCGGGGTAGTTCGCCGCTCTTCAACAACGACCTTGACGGAGGTG 18546
Qy 551 GATTACCTACCGTCCAACTTGAGGCTCTCTCAGGACCGGAGAAAGTTGGCGCACTGGC 610
Db 18545 TAGAACAACCCCTCCACGCTCTGTGGCCAGGTACACACCCCTGGTCACCGGTGCTCAAGTTG 18486
Qy 611 AAGCG 615
Db 18485 AAACG 18481

RESULT 13

US-09-426-436-1/c

; Sequence 1, Application US/09426436

; Patent No. 6225066

GENERAL INFORMATION:

APPLICANT: William R. Jacobs, Jr.
APPLICANT: Barry R. Bloom
APPLICANT: Graham F. Hatfull
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC
TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/426,436
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/705,557
FILING DATE:
APPLICATION NUMBER: US/08/057,531
FILING DATE:
APPLICATION NUMBER: 07/833,431

FILING DATE: February 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.
REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: 96700/238
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 286-0854 or 286-0082
TELEFAX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 52297
TYPE: nucleotide
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: phage genome sequence
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: not applicable.
ORIGINAL SOURCE:
ORGANISM: mycobacteriophage L5
STRAIN: not applicable
INDIVIDUAL ISOLATE: L5
DEVELOPMENTAL STAGE: not applicable
HAPLOTYPE: not applicable
CELL TYPE: not applicable
CELL LINE: not applicable
ORGANELLE: not applicable
IMMEDIATE SOURCE: mycobacteriophage L5 particles
POSITION IN GENOME: entire genome
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Hatfull and Sarkis
TITLE: DNA Sequence, Structure and Gene
TITLE: Expression of Mycobacteriophage L5:
TITLE: A Phage System for Mycobacterial
TITLE: Genetics
JOURNAL: Molecular Microbiology
VOLUME: 7
PAGES: 395-405
DATE: 1993.
US-09-426-436-1
Query Match 3.2%; Score 34.6; DB 3; Length 52297;
Best Local Similarity 49.2%; Pred. No. 4.7;
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Qy 431 TTGGCTAGACGACACCAAGAACCCATTGACGCTTCGCGATCCAACTATCGGTGATC 490
Db 18562 TTCCGCTGCTGAGGATGACGACCCCAAGACGATGTCGCTTCAGGGCTCGGTGGGAC 18503
Qy 491 AATTCGCGACACCGCGGACTCGGACTTATCCAGGCTTCGTGCGCACTTCG 550
Db 18502 AAGTAGCGAGCGCGGGGTAGTTCCTCCCGGCTCTTCAACACGACCTTGACGGAGGGTGC 18443
Qy 551 GATTACCTACGTCCTCACTTGGAGCCCTATCTCAGGACCGAGGAAAGTTGCGCAACTGGC 610
Db 18442 TAGAAACACCTCCACGCTGTGTGGCCAGGTACACACCTTGTCACCGGTGCTCAAGTTG 18383
Qy 611 AAGCG 615
Db 18382 AAACG 18378
RESULT 14
US-08-705-557-1/c
Sequence 1, Application US/08705557
Patent No. 630061
GENERAL INFORMATION:
APPLICANT: William R. Jacobs, Jr.
APPLICANT: Barry R. Bloom
APPLICANT: Graham F. Hatfull
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC
TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESS: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,557
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/057,531
FILING DATE:
APPLICATION NUMBER: 07/833,431
FILING DATE: February 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.
REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: 96700/238
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 52297
TYPE: nucleotide
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: phage genome sequence
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: not applicable.
ORIGINAL SOURCE:
ORGANISM: mycobacteriophage L5
STRAIN: not applicable
INDIVIDUAL ISOLATE: L5
DEVELOPMENTAL STAGE: not applicable
HAPLOTYPE: not applicable
CELL TYPE: not applicable
CELL LINE: not applicable
ORGANELLE: not applicable
IMMEDIATE SOURCE: mycobacteriophage L5 particles
POSITION IN GENOME: entire genome
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Hatfull and Sarkis
TITLE: DNA Sequence, Structure and Gene
TITLE: Expression of Mycobacteriophage L5:
TITLE: A Phage System for Mycobacterial
TITLE: Genetics
JOURNAL: Molecular Microbiology
VOLUME: 7
PAGES: 395-405
DATE: 1993.
US-09-426-436-1

```
; PAGES: 395-405
; DATE: 1993
; US-08-705-557-1

Query Match
Best Local Similarity 3.2%; Score 34.6; DB 4; Length 52297;
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 431 TTCGCTAGACACACCAAGAACCCATGAGCTTCGCGATCCAACTATCGGTGTATC 490
Db |||||
18562 TTCGGTCTGTTGAGGATGACGACCCCAAGACGATGTCTCGTTCAGGGCTCGGTGGAC 18503
QY 491 AATTCGCGACACCGCGGACTTCGGGACTTATTCAGGCTTCGTGTCGACCTTCG 550
Db |||||
18502 AAGTAGCGAGCGCGGGTAGTTCGCGGCTTTCACACGACCTTGACGGGAGGTGCG 18443
QY 551 GATTACCTACCGTCCAACTTGAGGCTATCTCAGGACCGAGAAAGTTGGCGCACTGGC 610
Db |||||
18442 TAGAACACACCTCCAGCTCTGTGCGCAGGTACACACCTGTACCCGTCGTCAAGTTG 18383
QY 611 AAGCG 615
Db |||||
18382 AACG 18378

RESULT 15
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (330001)..(345000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (345001)..(360000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (360001)..(375000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (375001)..(390000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (390001)..(405000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (405001)..(420000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (420001)..(435000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (450001)..(465000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (465001)..(480000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (480001)..(495000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (495001)..(510000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
```

```
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (870001)..(885000)
```

```
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
```

Query Match 3.1%; Score 33.2; DB 4; Length 1230025;
Best Local Similarity 56.4%; Pred. No. 69;
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```
Qy 929 ACGACGAGAACGTTCCCAAACTAGAGGGAATCGGAAATTTGGACCCCATGGTTGGCT 988
Db 751086 ATGACATAGTAATATCCGATATCAAAAGTTATCCAGACATTTGGTCATTTATTAGCAC 751027
Qy 989 GCGAAATAAACAAGGGCGCTGACCCAGGATAGTGAACCTTTGGAACTGGTT 1038
Db 751026 GCCATTACGCAAGTCAGAGAGATATGTTTTTATATTGACAGAGTGGCT 750977
```

Search completed: March 8, 2004, 01:21:58
Job time : 138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 8, 2004, 00:23:29 ; Search time 468 seconds
(without alignments)
8487.118 Million cell updates/sec

Title: US-09-993-777-1_COPY_9874_10959
Perfect score: 1086
Sequence: 1 ATGGCATCCTACTTGGAC.....CCCCGACTCAATAAATG 1086

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1086	100.0	13473	10	US-09-994-064-1
2	1084.4	99.9	3605	9	US-09-881-457A-3
3	1084.4	99.9	18913	10	US-09-994-064-59
4	343	31.6	534	10	US-09-994-064-18
5	38.2	3.5	777	12	US-10-282-122A-26478
6	38.2	3.5	780	12	US-10-282-122A-28211
7	37.6	3.5	6310	12	US-10-221-613-230
8	36.4	3.4	9025608	14	US-10-156-761-1
9	36	3.3	1049	14	US-10-123-155-358
10	36	3.3	1049	14	US-10-146-731-358
11	36	3.3	1049	14	US-10-140-472-358
12	36	3.3	1049	14	US-10-141-761-358
13	36	3.3	1049	14	US-10-142-885-358
14	36	3.3	1049	14	US-10-158-790-358
15	36	3.3	1049	15	US-10-137-871-358

C 16	36	3.3	1049	15	US-10-140-923-358	Sequence 358, App
C 17	36	3.3	1049	15	US-10-141-756-358	Sequence 358, App
C 18	36	3.3	1049	15	US-10-141-759-358	Sequence 358, App
C 19	36	3.3	1049	15	US-10-140-806-358	Sequence 358, App
C 20	36	3.3	1049	15	US-10-140-864-358	Sequence 358, App
C 21	35.6	3.3	1350	14	US-10-156-761-6893	Sequence 6893, App
C 22	35.2	3.2	19718	12	US-10-158-844-99	Sequence 99, Appl
C 23	35	3.2	653	14	US-10-184-644-402	Sequence 402, App
C 24	35	3.2	653	14	US-10-184-634-402	Sequence 402, App
C 25	35	3.2	1263	15	US-10-369-493-47306	Sequence 47306, A
C 26	35	3.2	1284	14	US-10-156-761-6164	Sequence 6164, Ap
C 27	35	3.2	9025608	14	US-10-156-761-1	Sequence 1, Appli
C 28	34.6	3.2	475	13	US-10-040-739-899	Sequence 899, App
C 29	34.6	3.2	493	10	US-09-918-995-21276	Sequence 21276, A
C 30	34.6	3.2	1620	15	US-10-094-749-1549	Sequence 1549, Ap
C 31	34.4	3.2	1563	14	US-10-156-761-3340	Sequence 3340, Ap
C 32	34.2	3.1	802	14	US-10-184-644-312	Sequence 312, App
C 33	34.2	3.1	802	14	US-10-184-634-312	Sequence 312, App
C 34	34	3.1	2855	9	US-09-764-970-594	Sequence 594, App
C 35	34	3.1	2855	14	US-10-125-540-594	Sequence 594, App
C 36	33.8	3.1	500	14	US-10-063-685-84	Sequence 84, Appl
C 37	33.4	3.1	1581	15	US-10-369-493-40657	Sequence 40657, A
C 38	33.2	3.1	498	11	US-09-864-408A-3455	Sequence 3455, Ap
C 39	33.2	3.1	777	14	US-10-022-832-1	Sequence 1, Appli
C 40	33.2	3.1	777	15	US-10-312-273-56	Sequence 56, Appl
C 41	33.2	3.1	1296	15	US-10-260-238-92	Sequence 92, Appl
C 42	33.2	3.1	1442	12	US-10-425-114-27719	Sequence 27719, A
C 43	33.2	3.1	1230025	15	US-10-289-762-1	Sequence 1, Appli
C 44	32.8	3.0	1215	14	US-10-156-761-6851	Sequence 6851, Ap
C 45	32.6	3.0	2010	14	US-10-431-304-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-994-064-1

; Sequence 1, Application US/09994064

; Publication No. US20030082788A1

; GENERAL INFORMATION:

; APPLICANT: Wild, Martha A.

; APPLICANT: Cochran, Mark D.

; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/994,064

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/468,190

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 39116-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:


```
; ORGANISM: Infectious Laryngotracheitis Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (585)..(1889)
; OTHER INFORMATION: ILTV glycoprotein D
; NAME/KEY: CDS
; LOCATION: (1997)..(3085)
; OTHER INFORMATION: ILTV glycoprotein I
US-09-881-457A-3

Query Match          99.9%; Score 1084.4; DB 9; Length 3605;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCGATCGCTACTTGGAACTCTGGCTCTCCCTTGGCGGACGCTCGCACCCCTTCGGCGCG 60
Db 1997 ATGCGATCGCTACTTGGAACTCTGGCTCTCCCTTGGCGGACGCTCGCACCCCTTCGGCGCG 2056

QY 61 ATGGGAATCGTGATCATCGGAATCATCGTCTCGGCCAGGATTGACGACGATCACATCGTG 120
Db 2057 ATGGGAATCGTGATCATCGGAATCATCGTCTCGGCCAGGATTGACGACGATCACATCGTG 2116

QY 121 ATCGTCGGCGCTCGCCCGGAGCTACAATTCAACTCGAGCTATTTTTCATGCTGCGCCAG 180
Db 2117 ATCGTCGGCGCTCGCCCGGAGCTACAATTCAACTCGAGCTATTTTTCATGCTGCGCCAG 2176

QY 181 AGACCCCAAAACCCCTACTCAGGAACCGTCGCGTCTCGGTCCTGATATAACAAC 240
Db 2177 AGACCCCAAAACCCCTACTCAGGAACCGTCGCGTCTCGGTCCTGATATAACAAC 2236

QY 241 CAGTGCTACAGGAACCTTAGCGAGGAGCGTTTGAATAATGCACTCATGCTGCTTCT 300
Db 2237 CAGTGCTACAGGAACCTTAGCGAGGAGCGTTTGAATAATGCACTCATGCTGCTTCT 2296

QY 301 GTTTTGTGCGGTGTAAGTACCGGAGTACAGTCTCGCGCTCGACAGACTTAACCGA 360
Db 2297 GTTTTGTGCGGTGTAAGTACCGGAGTACAGTCTCGCGCTCGACAGACTTAACCGA 2356

QY 361 CTCCACACCGCTTAAAGTCTACTATACGAATCTCTGCGGACGACGCGGATGTT 420
Db 2357 CTCCACACCGCTTAAAGTCTACTATACGAATCTCTGCGGACGACGCGGATGTT 2416

QY 421 TAGCTAATTGTTGCGGTAGACGACACCAAGAACCCATTGAGTCTTCCGATCCAACTA 480
Db 2417 TAGCTAATTGTTGCGGTAGACGACACCAAGAACCCATTGAGTCTTCCGATCCAACTA 2476

QY 481 TCGGTGTATCAATTCGCGAACACCGCGCGACTCTGCGGACTCTATTTCCAAAGGCTTCGT 540
Db 2477 TCGGTGTATCAATTCGCGAACACCGCGCGACTCTGCGGACTCTATTTCCAAAGGCTTCGT 2536

QY 541 CGCACCTTCGGATTACTACCGTCCAACTTGAGGCTATCTCAGGACCGAGGAAGTTGG 600
Db 2537 CGCACCTTCGGATTACTACCGTCCAACTTGAGGCTATCTCAGGACCGAGGAAGTTGG 2596

QY 601 CGCAACTGGCAAGCTACGTTGCGACGAGGCGACGACGACGAGCGCGGAGCGCAAC 660
Db 2597 CGCAACTGGCAAGCTACGTTGCGACGAGGCGACGACGACGAGCGCGGAGCGCAAC 2656

QY 661 CCGACGCGGCTACTGCAACACGCGCTCGCAACTTGAAGCGGAACACTTTACCTTTCC 720
Db 2657 CCGACGCGGCTACTGCAACACGCGCTCGCAACTTGAAGCGGAACACTTTACCTTTCC 2716

QY 721 TGGCTAGAAATGGCGTGATCATTTAGAACCGACACCGCGAAAACGAAATTCAAACGTT 780
Db 2717 TGGCTAGAAATGGCGTGATCATTTAGAACCGACACCGCGAAAACGAAATTCAAACGTT 2776

QY 781 ACTGTCGCTCGGGAACATGAGCCCTACGTAATTTGGGTAACTGCGTGGCTCGCGTCG 840
Db 2777 ACTGTCGCTCGGGAACATGAGCCCTACGTAATTTGGGTAACTGCGTGGCTCGCGTCG 2836

QY 841 AGCGCAACGATCGGCTCTGCTATTGTAATTTTCATCGTCAACGAAACATGTGCAACCG 900
Db 2837 AGCGCAACGATCGGCTCTGCTATTGTAATTTTCATCGTCAACGAAACATGTGCAACCG 2896
```

RESULT 3

```
US-09-994-064-59
; Sequence 59, Application US/09994064
; Publication No. US20030082788A1
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA: US/09/994,064
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/468,190
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18913 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: N
; ANTI-SENSE: N
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 697..1533
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: complement (1900..2784)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
```

12871	Db	CCTCACACCCGGTTTAAAGCTCACTATACGAATCTCTCGTCCGAACGACAGCGGATGTTCT	12930
421	Qy	TACGTAATTGTTTCGGCTAGACACACCAAGAAACCCATTGACGTCTTTCGGATCCAACTA	480
12931	Db	TACGTAATTGTTTCGGCTAGACACCAAGAAACCCATTGACGTCTTTCGGATCCAACTA	12990
481	Qy	TCGGTGTATCAATTTCGGAACAACCGCCGCGACTTCGCGACTCTATTCGAAGCTTCGTGT	540
12991	Db	TCGGTGTATCAATTTCGGAACAACCGCCGCGACTTCGCGACTCTATTCGAAGCTTCGTGT	13050
541	Qy	CGCACTTCGGAATCACTACGTTCCAACTTGAGGCTATCTCAGACCGAGGAAGTTGG	600
13051	Db	CGCACTTCGGAATCACTACGTTCCAACTTGAGGCTATCTCAGACCGAGGAAGTTGG	13110
601	Qy	CGCACTTCGGAAGCTGAGTTGCCACGAGAGCCACGACACCGCGCGAGGCGACAAC	660
13111	Db	CGCACTTCGGAAGCTGAGTTGCCACGAGAGCCACGACACCGCGCGAGGCGACAAC	13170
661	Qy	CGAGCCCGTCACTGCAACGAGCGCTCCGAACTTGAGCGGAACAATTTACCTTTCCC	720
13171	Db	CGAGCCCGTCACTGCAACGAGCGCTCCGAACTTGAGCGGAACAATTTACCTTTCCC	13230
721	Qy	TGCGTAGAAATTCGGTGGATCATTTACGAACCGCACACCCGCAACGAAATTCAAACGTT	780
13231	Db	TGCGTAGAAATTCGGTGGATCATTTACGAACCGCACACCCGCAACGAAATTCAAACGTT	13290
781	Qy	ACTGTCGCTTCGGACAATAGACCTTACGTTAATTTGGGGTAAACGTGGCTGCCTCGTG	840
13291	Db	ACTGTCGCTTCGGACAATAGACCTTACGTTAATTTGGGGTAAACGTGGCTGCCTCGTG	13350
841	Qy	AGGCGAACGATCGGCTCGTCAATTTCCATTCGTCACGAGAAACATGTGCACCCCG	900
13351	Db	AGGCGAACGATCGGCTCGTCAATTTCCATTCGTCACGAGAAACATGTGCACCCCG	13410
901	Qy	CACCGAAATTAGACACGCTTCGCAAGACGAGAAACGTTCCAACTAGAAGGAA	960
13411	Db	CACCGAAATTAGACACGCTTCGCAAGACGAGAAACGTTCCAACTAGAAGGAA	13470
961	Qy	TCGCGAAATTTGGACCCCATGGTTCCGTCGGAATAAACAAGGGCGTGCACGATAGT	1020
13471	Db	TCGCGAAATTTGGACCCCATGGTTCCGTCGGAATAAACAAGGGCGTGCACGATAGT	13530
1021	Qy	GAACTTGTGGAACGTGGTTGCGAATGTTTAAACCGCTTCGCGCTAAGCTCGCCGACTCAATA	1080
13531	Db	GAACTTGTGGAACGTGGTTGCGAATGTTTAAACCGCTTCGCGCTAAGCTCGCCGACTCAATA	13590
1081	Qy	AAAATG 1086	
13591	Db	AAAATG 13596	

RESULT 4

US-09-994-064-18/c
; Sequence 18, Application US/09994064
; Publication No. US20030082788A1

GENERAL INFORMATION: Martha A.
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSER: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

/ LOCATION: complement (2916..3605)
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3694..5124
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 5210..7081
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 7245..8123
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 8333..11290
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 11098..12402
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 13792..15291
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 15298..16080
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 16129..17013
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: complement (17380..18216)
/ OTHER INFORMATION:
/
/ US-09-994-064-59
/
Query Match          99.9%; Score 1084.4; DB 10; Length 18913;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1085; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGCATCGCTACTTGGAACTCTGGCTCTCTCTCTGCGCGAGCGCTGCAACCTTCGGGCGG 60
Db      12511  ATGCATCGCTACTTGGAACTCTGGCTCTCTCTCTGCGCGAGCGCTGCAACCTTCGGGCGG 12570

QY      61  ATGGGAATCGTATCACTGGAATCACTGCTCTGCGCAGGATGACGAGATCAATCGTG 120
Db      12571  ATGGGAATCGTATCACTGGAATCACTGCTCTGCGCAGGATGACGAGATCAATCGTG 12630

QY      121  ATCGTCGCGCTCGCCCGGAGCTACAATTCACTGCAGCTATTTTTCATGCTGGCCAG 180
Db      12631  ATCGTCGCGCTCGCCCGGAGCTACAATTCACTGCAGCTATTTTTCATGCTGGCCAG 12690

QY      181  AGACCCCAAAACCTTACTCAGAAACCGTCCGCGTCGGGTTTCGGTCTGTATATAACAAAC 240
Db      12691  AGACCCCAAAACCTTACTCAGAAACCGTCCGCGTCGGGTTTCGGTCTGTATATAACAAAC 12750

QY      241  CAGTGCTACAGGAACTTAGCGAGGAGCGCTTTGAAAATGCACTCATCGATCGTCTTCT 300
Db      12751  CAGTGCTACAGGAACTTAGCGAGGAGCGCTTTGAAAATGCACTCATCGATCGTCTTCT 12810

QY      301  GTTTTGTGGCTGTAAAGTACCGAGTACAGTTCTCCGCTCGAAACAGACTAACCGGA 360
Db      12811  GTTTTGTGGCTGTAAAGTACCGAGTACAGTTCTCCGCTCGAAACAGACTAACCGGA 12870

QY      361  CCTCCACACCGCTTTAAGTCACTATACGAAATCTCTGCTCCGAAACGACGAGCGGATGTC 420

```

SOFTWARE: PatentIn Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/994,064
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/469,190
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 391116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..533
US-09-994-064-18

Query Match 31.6%; Score 343; DB 10; Length 534;
Best Local Similarity 100.0%; Pred. No. 4.1e-109;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 TTACGACCGACACCGCGCAACGAAATTCACACGTTACTGTCGCTCGGACCAATGAG 803
DB 534 TTACGACCGACACCGCGCAACGAAATTCACACGTTACTGTCGCTCGGACCAATGAG 475
QY 804 CCCTAGCGTAATTTGGGGTAACCGTGGCTGCGGTGAGCGCAACCGATCGGCTGTCAT 863
DB 474 CCCTAGCGTAATTTGGGGTAACCGTGGCTGCGGTGAGCGCAACCGATCGGCTGTCAT 415
QY 864 TGTAATTTCCATCGTCACGAGAACATGTGCAACCGCGCCGCAAAATAGACACGGTCTC 923
DB 414 TGTAATTTCCATCGTCACGAGAACATGTGCAACCGCGCCGCAAAATAGACACGGTCTC 355
QY 924 GCAAGACGACGAGAACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCATGGT 983
DB 354 GCAAGACGACGAGAACGTTCCCAACTAGAGGGAATCGGAAATTTGGACCCATGGT 295
QY 984 TGGCTGCGGAATAAACAAGCGCGCTGACCGAGGATAGTGAATTTGGAACTGTTGGCAT 1043
DB 294 TGGCTGCGGAATAAACAAGCGCGCTGACCGAGGATAGTGAATTTGGAACTGTTGGCAT 235
QY 1044 TGTTAACCCTGTCGCTAAGCTCGCGCCGCACTCAATAAAATG 1086
DB 234 TGTTAACCCTGTCGCTAAGCTCGCGCCGCACTCAATAAAATG 192

RESULT 5
US-10-282-122A-26478/c
Sequence 26478, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 26478
LENGTH: 777
TYPE: DNA
ORGANISM: Mycobacterium bovis
US-10-282-122A-26478

Query Match 3.5%; Score 38.2; DB 12; Length 777;
Best Local Similarity 58.3%; Pred. No. 0.036;
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 29 TCCTTGGCGGACGCTCGCACCTTCGCGCGATGGGAATCGTGATCACTGGAATCACG 88
DB 544 TCCACGCGCCACGCTGCGACCTTCGTTGGCGGAGCAACGGTGATCACCCGTCACG 485
QY 89 TCTCCGCGAGATTCAGCAGCATCACATCGTCTCGGCGCTTCGCGCCGCAAGC 143
DB 484 TCGGCGCGGACGCGGTCTCCAGCGCGAGCAACCGCGCGCAAGCAGCGCGC 430

RESULT 6
US-10-282-122A-28211/c
Sequence 28211, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23

;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/230,335
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/230,347
;; PRIOR FILING DATE: 2000-09-09
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 28211
;; LENGTH: 780
;; TYPE: DNA
;; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28211

Query Match 3.5%; Score 38.2; DB 12; Length 780;
Best Local Similarity 58.3%; Pred. No. 0.036;
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 29 TCCTTGGCGGACGCTGCACCTTCGGCGGATGGGAATCGTGTGATCTGGAATCAG 88
|||
Db 544 TCCACGCGCCACGCTGCGACCTTCGGTGGCGGACGAGTGTATCACCCTCGTCAGCG 485
|||

Qy 89 TCTCGCGCAGGATTGACGACGATCACATCGTGTGTCGGCGCTTCGCCCGAAGC 143
|||
Db 484 TCGCCCCCAGGACGCGTGTCCAGCGGAGCGGAACCGCGCCCAAGCACCGCGCGC 430
|||

RESULT 7
US-10-221-613-230/c
; Sequence 230, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPERBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 230
; LENGTH: 6310
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-230

Query Match 3.5%; Score 37.6; DB 12; Length 6310;
Best Local Similarity 49.0%; Pred. No. 0.18;

Matches 100; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 611 AAGCGTACGTTGCCACGAGCGCCAGGACGAGCGGAGCGGAGCGCAACCCGACGCCG 670
|||
Db 1293 AAACGCCCGGACCCCGGAATACAAACGAGCGCGCGCTCCCGGCGCTCCCGGCGCA 1234
|||

Qy 671 TCACTGCAACGAGCGCTCCGAACTTGAAGCGGAACTTTTACCTTTTCCCTGGCTAGAAA 730
|||
Db 1233 ACCCGGATCCCGGACCCCGAACCAGACACTCTCGAACTCCGACTCCCGCGGCAAA 1174
|||

Qy 731 ATGGCGTGGATCATTACGACGACCGCGCAACGAAATTCAAAGTTTACTGTCCGTC 790
|||
Db 1173 ATAACTAACCGACTACGAAACCGCGCGCGGCGGAAACGAAACGAAACACCGTAGCTTC 1114
|||

Qy 791 TCGGACAAATGAGCGCTTACGCTAA 814
|||
Db 1113 GCCCGCAAAACGCCCTCCGCAA 1090
|||

RESULT 8
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 3.4%; Score 36.4; DB 14; Length 9025608;
Best Local Similarity 49.0%; Pred. No. 21;
Matches 97; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 494 TCGCGAACACCGCGGACTCCGAGCTCTATTCCAAGCTTCTGTGTGCGACCTTCGGAT 553
|||
Db 7090210 TCGGGCGCGCGCGGAGCTCCGCGCGCGCTCTGGCGCGCGCGCGCGCTCCGGG 7090151
|||

Qy 554 TACCTACGTTCCAACTTGAGGCTATCTCAGGACCGAGGAAAGTTGGCGCAACTGGCAG 613
|||
Db 7090150 CGCGCGCGCGCGCGGAGCTCCCTCCCGCGCGCGGATGGCGCGCTTCGCGAGG 7090091
|||

Qy 614 CGTACGTTGCCAGGAGCGCACGACGAGCGGCGGCGGCAACCCCGAGCGCGCTCA 673
|||
Db 7090090 GGCCTCCGTCGCCCGGAGCACCGGACCGCCACCGCGGCGGAGGACCTTGACCCCGGCGC 7090031
|||

Qy 674 CTGCAACGAGCGCTCCG 691
|||
Db 7090030 ACCGGCGCGCGCGCGCGG 7090013
|||

RESULT 9
US-10-123-155-358/c

```

/ Sequence 358, Application US/10123155
/ Publication NO. US20030068794A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filwaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3330R1C30
/ CURRENT APPLICATION NUMBER: US/10/123,155
/ CURRENT FILING DATE: 2002-04-15
/ Prior Application removed - See Palm or File Wrapper
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 358
/ LENGTH: 1049
/ TYPE: PRT
/ ORGANISM: Homo Sapien
/ US-10-123-155-358

```

Query Match	3.3%; Score 36; DB 14; Length 1049;
Best Local Similarity	4.5%; Pred. No. 0.25;
Matches	35; Conservative 215; Mismatches 529; Indels 1; Gaps 1;
QY	270 CTTTGAAATGCACTCATCGATCGTCTTCCTGTTTTGCGGCTGTAAGTGACCGACTA 329
Db	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db	818 C.CBGABAT.R..AB.ADNBWB.WBTHGAG.YYND..RX.NN.BN..SAM..K.. 759
QY	330 CAGGTTCTCGGCTCGAAGACTAACCGGACCTCCACACCGCTTAACTCACTATACG 389
Db	758 MNSS.H.RY...TH...RMA.SY..NNM...NM.SYSGNS.Y..BAA..NDS.H.A..NM 699
QY	390 AATACCTC-GTCCGAACGACAGGGATGTTCTACGTAATGTTGCGCTAGACGACCA 448
Db	698 .G..MWS.SM.CNMT.S.NM.N..KCH.BCS...S.SNMS.H....M..NM....RYNH 639
QY	449 AAGAACCATTGACGCTCTTCGCTATCCAACTATCGGTGTATCAATTCGCGACACCGCG 508
Db	638 C.YW.BH.DNCY...AY.S.S.KAYSASS.HHNFK.M..BM.NMA.N.KDA.C.S..R 579
QY	509 CGACTCGCGGACTCTATTCCAAGGCTTCGTGTCGCACCTTCGGATTACCTACCGTCCAAC 568
Db	578 DNSS.H.B..MD...TASD..H.YNNS.H.RY..T.....SCN.A.S..NCS.N.GM. 519
QY	569 TTGAGGCCATCTCAGGACCGAGGAAAGTTGGGCGCACTGGCAAGCGTACGTGCCACGG 628
Db	518 .S.D..HSSMB...SNMS.H.A.CPMRGS.NBSK.ST.MNM.VGSYTRMRY.RD...B 459
QY	629 AGGCACGACGACGACGCGCGAGCGACAACCCCGACGCCGCTCACTGCAACACGCGCT 688
Db	458 ...RS.BSAYTNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.M..NAC.H.B.. 399
QY	689 CCGAATCTGAACGGGAACACTTTACCTTCCCTCGGCTAGAAAATGCGGTGGATCATTTACG 748
Db	398 N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYRB....N.S.H....S.. 339
QY	749 AACCGACACCGCGAACAAGAAATTCAAACGTTACTGCTGTCGGGCAATGACGCCCTA 808
Db	338 .D..MTHC..MT..N.S.H....MN.NM.WY..BD..SNSD.Y.BM..A.TH.TWB.... 279

```

Qy 809 CGCTAATTGGGGTAAACCGTGGCGTCCGTGTGAGCGCAACGATCGCGCTCGTCAITGTGTA 868
Db 278 .SNNMG.TG...TNRGY.GNCS.H....N.NN.HH....MT.KNNR.R..A.AS..BA.B 219
Qy 869 TTTCCATGCTCACCAGAAACATGTGCACCCGCGACCGAAAATTAGACACGGTCTCGCAAG 928
Db 218 TAGNNHM.S.BM.MA.N..THM..SRSBG.NYRRN.C.R.....NT..A.N.MY.S..N 159
Qy 929 ACGACGAAGAACGTTCCCAAACTAGAAGGGAATCGCGAAAATTGGACCCCATGTTGCGT 988
Db 158 NT..S.....S.C.....NCH.R.SM.RA.CS.SY.M...YM.GKNMSC....BGN 99
Qy 989 GCGAAATAACAAGGCGGTGACACGATAGTGAATCTGTGGAACTGGTTCGCAATGTTA 1048
Db 98 GY.H..B.DH.YD.ST.S.H..DN.A.A.NAANA..CC...A.DMHAGHB.BDNM.BH.A 39

RESULT 10
US-10-146-731-358/c
; Sequence 358, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: F3330R1C323
; CURRENT APPLICATION NUMBER: US/10/146,731
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-731-358

Query Match 3.3%; Score 36; DB 14; Length 1049;
Best Local Similarity 4.5%; Pred. No. 0.25;
Matches 35; Conservative 215; Mismatches 529; Indels 1; Gaps 1;

Qy 270 CTTTGAATAATGCATCATGATCGTCTCTCTCTGTTTGTGCGGTGTAAGTGACCGAGTA 329
Db 818 C.CEBGABAT.R..AB.ADNBWB.WBTHGAG..YND..KV.NN.BN...SAY..K.. 759
Qy 330 CAGTCTCGCGCTCGAAGACAGACTACCGGACCTCCACACCGGTTTAAGTCACTATACG 389
Db 758 MNSS.H.R.Y...TH...RMA.SY..NNM...NM.SYSGNS.Y..BAA..NDS.H.A..NM 699
Qy 390 AAATCTCTC-GTCCGAACGACAGCGGGATGTTCTACGTAATTTGTCGGCTAGACGACCA 448
Db 698 .G..MMWS.SM.CNMT.S.NM.N..KCH.BCS....S.SNMS.H....M.NM....RYNH 639
Qy 449 AAGAACCATTGACGCTCTTCGGCATCAACTATCGGTGTATCAATTCGGAACACCGCGC 508
Db 638 C.YW.BH.DNCY...AY.S.S.KAYSASS.HNHKK.M..BM.NMA.N.KDA.C.S..R 579
Qy 509 CGACTCGCGACTCTATTCCAAGGCTTCGTGTGCGACCTTCGGATTAACCTACGTCAC 568

```

578	DSNSS.H.B..MD.....TASD..H.VNNS.H.R.Y..T.....SCN.A.S..NCS.N.GM..	519
569	TTGAGGCGCTATCTCAGGACCGCAGGAAAGTTGGCGCAACTGCGCAAGCGTACGTTGCCACGG	628
518	S.D..HSSMB...SNMS.H.A.CMRG.S.NBSK.ST.MNM.VGSTRHRY.RD....B	459
629	AGGCCAGCAGCAGCAGCGCGGAGCGCAACCCGAGCGCCGTCACCTGCAACAGCGCCT	688
458	...RS.BSAVTSNG.CB.SSHCS.S.MVBS.H.BM.YM..M.KS.NT.M..NAC.H.B..	399
689	CGGAACCTTGAAGCGGAACACTTTACCTTTCCTCGGTAGAAAATGGCGTGATCAITACG	748
398	N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYRB....N.S.H....S...	339
749	AACCGACACCGCGCAACGAAATTCAAACGTTACTGTCCTCGTCGGACAATGAGCCCTA	808
338	D..MTEC..MT..N.S.H....NM.NM.WY..BD..SNSD.Y.BM..A.TH.TNB....	279
809	CGCTAATTTGGGTAAACGTCGTGGCTCCGTCGTGAGCGCAACGATCGGCGCTCGTCATGTAA	868
278	SNNMG.TG...TNRGY.GNCS.H....N.NN.HH...MT.KNKR.R..A.AS..BA.B	219
869	TTTCCATCGTCACCGAGAAACATGTGCACCGCGCAGCAAGAAATTAGACACGCTCTCGAAG	928
218	TABNNHM.S.BM.NA.N..THM..SRBGR.NYRRGN.C.R....NT..A.N.MY.S..N	159
929	ACGACGAAGAACGTTCCCAAACTAGAAGGAAATCGCGAAAATTTGGACCCATGTTGCGT	988
158	NT..S....S..C.....NCH.R.SM.RA.CS.SY.M..YM.GKNWSC....BGN	99
989	CGCAATAAACAAGGGCGCTACACAGGATGATGAACTTTGTGGAACTCGTTGGCATGTTTA	1048
98	GY.H..B.DH.YD.ST.S.H...DN.A.A.NAANA..CC...A.DMAGHB.BNNM.BH.A	39

RESULT 11

US-10-140-472-358/c
; Sequence 358, Application US/10140472
; Publication No. US20030138888a1

GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.
 / APPLICANT: Beresini, Maureen
 / APPLICANT: DeForge, Laura
 / APPLICANT: Desnoyers, Luc
 / APPLICANT: Filvaroff, Ellen
 / APPLICANT: Gao, Wei-Qiang
 / APPLICANT: Gerritsen, Mary E.
 / APPLICANT: Goddard, Audrey
 / APPLICANT: Godowski, Paul J.
 / APPLICANT: Gurney, Austin L.
 / APPLICANT: Sherwood, Steven
 / APPLICANT: Smith, Victoria
 / APPLICANT: Stewart, Timothy A.

APPLICANT: Tamas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3330R1C16

1. CURRENT APPLICATION

1. CURRENT FILING

1. Prior Application

; NUMBER OF SEQ ID NOS: 5

; SEQ ID NO 358

; LENGTH: 1049

```

; TYPE: PRT

```

; ORGANISM: Homo Sapien

US-10-140-472-358

Matches	35; Conservative	215; Mismatches	529; Indels	1; Gaps
QY	270	CTTTGAAATTCACATCATCGATCGCTCTTCTGCTTTTTCGCGCTGTAAGTGCACGAGTA	329	
Db	818	C.BGABHAT.R..AB.ADNBWB.WBTHGAG..YND.D...KM.NN.BN...SAM..K..	759	
QY	330	CAGGTTCTCGGCTCGAAGACTAACCGGACTCCACACCGTTTAACTCTACTATACG	389	
Db	758	MNSS.H.RY...TH...RMA.SY..NNM...NM.SYGN.S.Y..BAA..NDS.H.A..NM	699	
QY	390	AAATCTCTC-GTCCGACGACGCGGATGTTCTAGTAATTTCTCGGCTAGACGACCA	448	
Db	698	G..MMWS.SV.CNWT.S.NM.N..KCH.BCS...S.SNWS.H...M.NM...RYNH	639	
QY	449	AAGAACCATTACGCTTTTCGGATCCAACTATCGGTATCAATTCGCGAAGACGCGCG	508	
Db	638	C.YW.BH.DNCY...AY.S.S.KAYSASS.HHNKK.M..BM.NMA.N.KDA.C.S..R	579	
QY	509	CGACTCGGAGCTTATTCCAAGGTTTCGTGTGCGACCTTCGATTACCTACGTC	568	
Db	578	DSNSS.H.B..MD...TASD..H.YNNS.H.RY..T.....SCN.A.S..NCS.N.GM.	519	
QY	569	TTGAGGCTTATCTCAGGACCGGAGAAAGTTGGCGAACTGGCAAGCGTACGTTGCCACG	628	
Db	518	S.D..HSNB...SNMS.H.A.CMRGS.NESK.ST.MNM.YGSYTRHRY.RD....B	459	
QY	629	AGCCACGACGACGCGCGGAGGCGAACAACCGGACGCCCTCACTGCAACAGCGCCT	688	
Db	458	..RS.BSAYTNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.M..NAC.H.B..	399	
QY	689	CGAATCTGAAGGGAACACTTACTTTCCTTCGCTAGAAAATGCGTGGATCATACG	748	
Db	398	N..ND..S.N.SM..M.BCY.Y..M.SM.SS.T.S.NKSTYRB...N.S.H...S...	339	
QY	749	AACGACACCCGACGAAATTCAAAGTTACTGCTCGCTCTCGGACAAATGAGCCCTA	808	
Db	338	D..MTHC..MT..N.S.H...MN.NM.WY..BD..SNSD.Y.EM..A.TH.TNB....	279	
QY	809	CGCTAATTCGGGTAAACGCTGGTGGCTGCGTGGAGCGCAACGATCGGCCTCGTCA	868	
Db	278	..SNMG.TG...TNRGY.GNCS.H.....N.NM.HH...MT.KNNR.R..A.AS..BA	219	
QY	869	TTTCCATCGTCACGAAACATGTGCACCCCGCACCGAAATAGACACGGTCTCGAAG	928	
Db	218	TABNNHM.S.BM.MA.N..THM..SRBRG.NYRRGN.C.R...NT..A.N.MY.S..N	159	
QY	929	ACGACGAGACGTTCCCAACTAGAGGGAATCGGAAAATTTGGACCATGTTGCGT	988	
Db	158	NT..S...S...C.....NCH.R.SM.RA.CS.SY.M...YM.GKNMSC....BGN	99	
QY	989	GCGAATAACAAAGGCGGCTGACCGAGATGTAACCTTGTGGAACCTGGTTCGATTGTTA	1048	
Db	98	GY.H..B.DH.YD.ST.S.H..DN.A.A.NAANA..CC...A.DMHAGHB.BDNM.BH.A	39	

RESULT 12
 US-10-141-761-358/c
 ; Sequence 358, Application US/10141761
 ; Publication No. US20030148432A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.

RESULT 12

US-10-141-761-358/C

: Sequence 358, Application US/101411761

: Publication No. US20030148432A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E

; APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

Query Match	3.3%; Score 36; DB 15; Length 1049;
Best Local Similarity	4.5%; Pred. NO. 0.25;
Matches	Conservative 215; Mismatches 529; Indels 1; Gaps 1;
Qy	270 CTTTGAATAATTGCACCTCATCGATCGCTCTCTGTTTGTTCGGCTGTAAGACGTGACCGGAGTA 329
Db	818 C.CEGBABT.R.AB.ADNBWB.WETHGAG.YYDD...KM.NN.BN...SAM..K.. 759
Qy	330 CAGCTTTCGCGCTCGAACAGACTAACCGGACCTCCACACCGGTTTAAGCTTCACTATACG 389
Db	758 MNSS.H.RY..TH..RMA.SY..NNM..NM.SYSGNS.Y..BAA..NDS.H.A..NM 699
Qy	390 AAATCCTC-GTCCGAACGACAGCGGGATGTTCTACGTAATGTTTCGCTAGACGACACCA 448
Db	698 G..MMWS.SM.CNMT.S.NM.N..KCH.BCS...S..SNMS.H...M..NM...RYNH 639
Qy	449 AAGAACCCATTTGACGTCTTCGCGATCAAACTATCGGTGTATCAATTCGGAACACCGCCG 508
Db	638 C.YW.BH.DNCY...AY.S.S.KAYSASS.HNHKK.M..BM.NMA.N.KDA.C.S..R 579
Qy	509 CGACTCGCGACTTATTCCAGGCTTCGTGTCGCACCTTCGGAATTTACCTACGTCGTCGAC 568
Db	578 DSNSS.H.B..MD...TASD..H.YNNS.H.RY..T....SCN.A.S..NCS.N.GM.. 519
Qy	569 TTGAGGCTATCTCAGGACCGAGGAAAGTTGGCGCACTGCGAAGCGTACGTTGCCACCG 628
Db	518 .S.D..HSSMB...SNMS.H.A.CRMGGS.NBSK.ST.MNM.VGSYTRMRY.RD...B 459
Qy	629 AGGCGACGACGACCGGCGGAGCGCAACCCGAGCGCCGTCACCTGCACCAACGAGCGCCT 688
Db	458 ...RS.BSAYTNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.M..NAC.H.B.. 399
Qy	689 CCGAATTGAACGGGAACATTTACCTTTCCCTGGCTAGAAATGGCGTGATCATTTACG 748
Db	398 N..ND..S.N..SM..M.BRCY.Y.M.SM..SS.T.S.NKSTYRB...N.S.H....S.. 339
Qy	749 AACCGACACCGCAACGAAATTCAAACGTTACTGTCGCTCGGACAAATGAGCCCTA 808
Db	338 .D..MTHC..MT..N.S.H....MN.NM.WY..BD..SNSD.Y.BM..A.TH.TNB.... 279
Qy	809 CGCTAAATTGGGGTAAACGTCGCTGCCGTCGTGAGCGCAACGATCGGCTCGTCATTGTAA 868
Db	278 .SNNVG.TG...TNRGY.GNCS.H....N.NN.HH...MT.KNMR.R..A.AS..BA.B 219
Qy	869 TTTCCATGTCACAGAAACATGTGACCCCGCACCGGCAACAAAATTAGACACCGTCTCGAAG 928
Db	218 TABNNHM.S.BM.NA.N..THM..SRSBG.NYRGRN.C.R....NT..A.N.MY.S..N 159
Qy	929 ACGACGAAGAACGTTCCCAACTAGAAGGGAATCGGAAATTTGGACCATCGTTGCGT 988
Db	158 NT..S....S..C.....NCH.R.SM.RA.CS.SY.M..YM.GKNVSC....BGN 99
Qy	989 GCGAAATAACAAGGCGCTGACACGAGATGTAACCTTTGTGGAACTCGTTGCGATGTTTA 1048
Db	98 GY.H..B.DH.YD..ST.S.H..DN.A.A.NAANA..CC...A.DMHAHB..BDNM.BH.A 39

Search completed: March 8, 2004, 02:39:41
Job time : 487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 22:42:19 ; Search time 3277 Seconds
(without alignments)
9896.347 Million cell updates/sec

Title: US-09-993-777-1_COPY_9874_10959
Perfect score: 1086
Sequence: 1 ATGGCATCGTACTTGGAAC.....CGCCGACTCAATAAAATG 1086

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43.2	4.0	493	10 BE777414	BE777414 MY-35-A-0
C 2	43.2	4.0	885	13 BX425603	BX425603 BX425603
C 3	42	3.9	681	9 AU076048	AU076048 AU076048
C 4	41.4	3.8	591	9 AU068783	AU068783 AU068783

5	41.4	3.8	765	28 AZ132997	AZ132997 OSJNB010
6	41.4	3.8	1053	29 CMS060TI	AL411100 T7 end of
7	41.2	3.8	453	14 CA704113	CA704113 wdkic.pk0
C 8	41	3.8	618	14 CB883239	CB883239 HQ01J24W
C 9	40.2	3.7	852	13 BX393687	BX393687 BX393687
C 10	39.8	3.7	1201	13 BX360624	BX360624 BX360624
C 11	39.4	3.6	328	29 CE179705	CE179705 tigr-gss-
C 12	39.4	3.6	551	29 CE554687	CE554687 tigr-gss-
C 13	38.6	3.6	577	14 CA255506	CA255506 SCEPFL418
C 14	38.6	3.6	717	14 CF877969	CF877969 trico077XP
15	38.6	3.6	718	14 CF865608	CF865608 trico002Xe
16	38.6	3.6	727	14 CF867578	CF867578 trico111Xc
17	38.6	3.6	735	14 CF866505	CF866505 trico066Xc
18	38.6	3.6	747	14 CF868489	CF868489 trico015Xc
19	38.6	3.6	750	14 CF866359	CF866359 trico006Xc
20	38.6	3.6	752	14 CF866952	CF866952 trico008Xc
21	38.6	3.6	772	14 CB906186	CB906186 trico077XP
22	38.6	3.6	775	14 CB895519	CB895519 trico002Xe
23	38.6	3.6	788	14 CB897819	CB897819 trico111Xc
24	38.6	3.6	791	14 CB896481	CB896481 trico006Xc
25	38.6	3.6	803	14 CB898569	CB898569 trico015Xc
26	38.6	3.6	805	14 CB896951	CB896951 trico008Xc
27	38.6	3.6	807	14 CB896324	CB896324 trico006Xc
28	38.6	3.6	1201	13 BX360624	BX360624 BX360624
C 29	38.2	3.5	575	12 B1352860	B1352860 GM21095.5
30	38.2	3.5	720	28 BZ896103	BZ896103 NARP6_012
31	38	3.5	732	14 CF868364	CF868364 trico14x1
32	38	3.5	788	14 CB898441	CB898441 trico14x1
33	37.8	3.5	865	14 CB899074	CB899074 trico17xc
34	37.8	3.5	865	14 CF868966	CF868966 trico17xc
C 35	37.8	3.5	925	29 CFS0091P	AL033013 prosophi1
C 36	37.6	3.5	693	14 CF882808	CF882808 trico087xf
37	37.6	3.5	749	14 CB909128	CB909128 trico087xf
38	37.6	3.5	1200	9 AL579901	AL579901 AL579901
39	37.4	3.4	1201	13 BX381961	BX381961 BX381961
C 40	37.2	3.4	591	14 CA181075	CA181075 SCAGST313
41	37	3.4	519	13 CA127859	CA127859 SCAGLR201
C 42	37	3.4	550	12 BM596053	BM596053 170006874
C 43	37	3.4	675	12 BM603983	BM603983 170006870
C 44	37	3.4	722	14 CF869283	CF869283 trico18xm
C 45	37	3.4	731	12 BM585009	BM585009 170006872

ALIGNMENTS

RESULT 1
BE777414/c
LOCUS BE777414 493 bp mRNA linear EST 20-SEP-2000
DEFINITION MY-35-A-04 PinfestansMY Phytophthora infestans cDNA, mRNA sequence.
ACCESSION BE777414
VERSION BE777414.1 GI:10231053
KEYWORDS EST.
SOURCE Phytophthora infestans (potato late blight agent)
ORGANISM Phytophthora infestans
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.

REFERENCE 1 (bases 1 to 493)

AUTHORS Kamoun,S., Hraber,P.T., Sobral,B.W.S., Nuss,D. and Govers,F.
TITLE Initial assessment of gene diversity for the oomycete pathogen
JOURNAL Phytophthora infestans based on expressed sequences
MEDLINE Fungal Genet. Biol. 28 (2), 94-106 (1999)
PUBMED 20056376

COMMENT 10587472

Contact: Govers F

Laboratory of Phytopathology

Wageningen University

Binnenhaven 9, P.O.Box 8025, 6700 BE, Wageningen, The Netherlands

Tel: 31 317 483 138

Fax: 31 317 483 412

Email: Francine.Govers@medew.fyto.wau.nl.

Location/Qualifiers

source 1. .493

FEATURES

/cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="22735.62"
 /dev_stage="flowering stage"
 /clone_lib="Rice panicle at flowering stage"
 /note="Organ: panicle; Rice cDNA from panicle at flowering stage"

ORIGIN

Query Match 3.9%; Score 42; DB 9; Length 681;
 Best Local Similarity 53.4%; Pred. No. 6.4;
 Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 560 CCGTCCAACTTGGGCTATCTCAGGACCGAGAAAGTTGGCGCAACTGGCGAGGCTAGG 619
 DB 354 CCATCCGGGTTGAGAGTGGTTTGAAGATGCTGACGTGGCGGCGGCTGGCCATGT 295
 QY 620 TTGCCAGGAGCCGACGACGACCGCGGAGCGGACACCGCCGCTCACTGCAA 679
 DB 294 CGGCGACCGAGGCCACCGCGGAGGAGCGGACGACGACGAGCGCGGCGGCT 235
 QY 680 CCAGCGGCTCCGAACTTGAAGCGGAACTTTACCTTTCCCTG 722
 DB 234 GCAGCGCCACCCACTTCGGCTGCTCTGNACTTCGCTG 192

RESULT 4

AU068783/c
 LOCUS AU068783
 DEFINITION AU068783 Rice callus Oryza sativa (japonica cultivar-group) cDNA clone C50531_102, mRNA sequence.
 ACCESSION AU068783
 VERSION AU068783.1 GI:5003634
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 591)
 AUTHORS Yamamoto, K. and Sasaki, T.
 TITLE Rice cDNA from callus (1998)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 Tel: 81-298-38-7441
 Fax: 81-298-38-7468
 Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
 PROEUCT = "RGP"

FEATURES

source
 1..591
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="C50531.102"
 /clone_lib="Rice callus"
 /note="Vector: pBluescript II SK+; Site 1: SalI; Site 2: NotI; cDNA prepared from rice callus mRNAs by using oligo(dT) as a primer and ligating to the SalI-NotI site of pBluescript II SK+ phagemid."

ORIGIN

Query Match 3.8%; Score 41.4; DB 9; Length 591;
 Best Local Similarity 53.4%; Pred. No. 8.7;
 Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 560 CCGTCCAACTTGGGCTATCTCAGGACCGAGAAAGTTGGCGCAACTGGCGAGGCTAGG 619
 DB 260 CCATCCGGGTTGAGAGTGGCTTGAAGATGCTGACGTGGCGGCGGCTGGCCATGT 201

RESULT 5

AZ132997
 LOCUS AZ132997
 DEFINITION OSJNB0108G22f CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica cultivar-group) genomic clone OSJNB0108G22f, genomic survey sequence.
 ACCESSION AZ132997
 VERSION AZ132997.1 GI:8211738
 KEYWORDS GSS.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 765)
 AUTHORS Wing, R.A. and Dean, R.A.
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
 JOURNAL Unpublished (1998)
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: GTAAACGACGCGCCAGTG
 Class: BAC ends
 High quality sequence start: 31
 High quality sequence stop: 580.

FEATURES

source
 1..765
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNB0108G22f"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice BAC Library (EcoRI)"
 /note="Vector: pBACindigo; Site 1: EcoRI; Site 2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

ORIGIN

Query Match 3.8%; Score 41.4; DB 28; Length 765;

	Best Local Similarity	53.48;	Pred. No.	9.6;	Mismatches	87;	Conservative	0;	Indels	76;	Gaps	0;
QY	560	CCGTCCAACTTGAGCGCCTATCTCAGGACCGAGGAAGAAGTTGGCGCAACTGGCAAGCGTACG	619									
Db	280	CCATCCGGGTTCAGAGCTGCCTTTGGAGATGTTGACGTGGCAGCGGCTGCCCATGT	339									
QY	620	TTCGCCAGGAGGCGACGACGACCAGCGCCGAGGCGCACACCCGACGCCCGTCACTGCAA	679									
Db	340	CGGCGCACCGAGCCCACCGCGCGGACGACGACGAGCGCGCGCCGACCGTTGAGCGCCT	399									
QY	680	CRAGCGCTCGAATTGACGCGAACACATTACCTTTCCCTG	722									
Db	400	GCAGCGCACCCCACTTCGCGTGCTCTCTGCACCTTCGCGCTG	442									

RESULT 6	CNS06QTI	1053 bp	DNA	linear	GSS 05-JUL-2001
LOCUS	T7 end of clone AWOAA009A05 of library AWOAA from strain CLIB 89 of Yarrowia lipolytica, genomic survey sequence.				
DEFINITION	Yarrowia lipolytica, genomic survey sequence.				
ACCESSION	AL411100				
VERSION	AL411100.1	GI:12180174			
KEYWORDS	GSS.				
SOURCE	Yarrowia lipolytica				
ORGANISM	Yarrowia lipolytica				
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Dipodascaceae; Yarrowia.				
AUTHORS	1 (bases 1 to 1053) Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,B., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Ozler-Kalogeropoulos,O., Potier,S., Saurin,W., Tekalia,F., Toffano-Nioche,C., Weslowski-Louvel,M., Wincker,P. and Weissenbach,J.				
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies				
JOURNAL	FEMS Lett. 487 (1), 3-12 (2000)				
MEDLINE	FEMS Lett. 487 (1), 3-12 (2000)				
PUBMED	20584711				
FEATURES	1..1053				
AUTHORS	2 (bases 1 to 1053) Casaregola,S., Neuveglise,C., Lepingle,A., Bon,E., Feynerol,C., Artiguenave,F., Wincker,P. and Gaillardin,C.				
TITLE	Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia lipolytica				
JOURNAL	FEMS Lett. 487 (1), 95-100 (2000)				
MEDLINE	FEMS Lett. 487 (1), 95-100 (2000)				
PUBMED	20584727				
REFERENCE	1152892				
AUTHORS	3 (bases 1 to 1053) Genoscope.				
TITLE	Direct Submision				
JOURNAL	Submitted (07-SP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. [E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr]				
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces Hansenii var. Hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.				
FEATURES	Location/Qualifiers				
source	1..1053 /organism="Yarrowia lipolytica" /mol_type="genomic DNA" /strain="CLIB 89" /db_xref="taxon:4952" /clone="AWOAA009A05" /clone_lib="AWOAA" /note="end : 77"				

863 TTGTAATTTTCATCGTCCAGAAACCATGTGCACCCTGGCAGCAATATTAGACACGGTCT 922
606 AMMMMAAMMMKKTKTWAKKMMMMMMMMMMMMMMMMMMGAAGAAGMMMMMMMMMMAGM 547
923 CGCAAGACGACGAGAACGTTGCCCAAACACTAGAACGGAATCGCGAAAA 969
546 MGVMMGMVAMGVMMGMGMVVMMMMMMMMMMMMMMMMMMMMMMMMMMMMHMM 500

```

RESULT 10
BX360624/c
LOCUS
DEFINITION BX360624 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSODI072YU05 3-PRIME, mRNA sequence.
ACCESSION BX360624
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li W.B., Gruber C., Jesses J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: sqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4982.r
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSODI072CF03NPl.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI072YL05"
/tissue-type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSORT 6 vector. Library was normalized."

FEATURES
source
Query Match 3.7%; Score 39.8; DB 13; Length 1201;
Best Local Similarity 11.3%; Pred. No. 30;
Matches 35; Conservative 149; Mismatches 125; Indels 0; Gaps 0;

Oy 478 CTATCGGTGATCAATTCGCAACACCGCCGCACTCGGACCTATTCCAGGCTTCG 537
Db 1186 BTTTSTGSBSHTTTSTBSTSSSSSSSBTTTTSSSSSBTSBSSSSSSSYSTTSS 1127
Oy 538 TGTGCACCTTCGGATTACCTACCGTCCAACCTTGAGGCTATCTCAGCACCGAGAAAGT 597
Db 1126 BBTSTBTBTGTTTCKSVVTCBTBTMSKTSTSCSYVTTCSTVTISCSBBCSTTSCSTT 1067
Oy 598 TGGCGCACTGGCAAGCGTAGTTCGCCAGGAGGCCACGACACGACCGCCGAGCGACA 657
Db 1066 CSTCTCSYSTSBBBCSBTTTTTSSSSSSSSSTTSTTBBSBSSSSSSSSSSSSSSSS 1007
Oy 658 ACCCGAGCCCGCTCACTGCAACACGAGCGCTCCGAACCTTGACCGGACACTTTACCTTT 717
Db 1006 SSSSSSSSSSSSBSTBTBSBSSSSSBTSSSVSSSVSSSSSSSSSSSSBSCSBSTSSS 947
Oy 718 CCCTGGCTAGAAAATGGCGTGGATCATTCAGAACCGACACCGCGAACAAGAAATTCAAAC 777
Db 946 SSSWSSSSSSSSSSSSSSSSSSSSSTSKWSSTASSSSSVSWTAATTTTAAWARLAM 887
Oy 778 GTTACTGTC 786

```

AUTHORS	Kirkness,B.P., Bafna,V., Halpern,A.L., Levy,S., Remington,K., Kusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.
TITLE	The dog genome: survey sequencing and comparative analysis
JOURNAL	Science 301 (5641), 1998-1903 (2003)
MEDLINE	22875432
PUBMED	14512627
COMMENT	Contact: Kirkness BP The Institute for Genomic Research Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-0200 Fax: 301-838-0208 Email: ekirknes@tigr.org Class: shotgun.
FEATURES	Location/Qualifiers 1..551 /organism="Canis familiaris" /mol_type="genomic DNA" /strain="Standard Poodle" /db_xref="taxon:9615" /clone_lib="Dog Library" /note="Site 1: BetXI; Libraries were prepared from peripheral blood"
ORIGIN	Query Match 3.6%; Score 39.4; DB 29; Length 551; Best Local Similarity 52.8%; Pred. No. 28; Matches 85; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY	454 CCCATGACGCTTCGGGATCCAACTACGGTATATCAATTGCGAACAACGCCGCGACT 513
Db	14 CACCTGGACCTCTCTCTGGAACCTGGTGACTCTCTTTCCAACTGGAACCTCACCTGGATC 73
QY	514 CCGGACGCTATTCCAGGCTTCGTGCGACCTTCGGATTACCTACCGTCCAACTTGAG 573
Db	74 TGGTGACTCTGGTCTCTCAGTTGGCTTCACCTGGACCTGGTGCTCTGTCTGTCCACACTTTG 133
QY	574 GCGTATCTCAGACCGAGGAAAGTTGGCGCAACTGGCAAGC 614
Db	134 ACCTAACTCAGATGCTGTGATCTTGGCCCCACCTAGCGCGC 174
RESULT 13.	
CA255506/c	
LOCUS	SCEPFL180E09.g 577 bp mRNA linear EST 26-SEP-2003
DEFINITION	5', mRNA sequence.
ACCESSION	CA255506
VERSION	CA255506.1 GI:35940463
KEYWORDS	EST.
SOURCE	Saccharum officinarum
ORGANISM	Saccharum officinarum
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.
AUTHORS	1 (bases 1 to 577)
TITLE	Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
JOURNAL	The libraries that made SUCEST
COMMENT	Genet. Mol. Biol. 24 (1-4), 1-7 (2001) Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br Plate: 180 row: E column: 09 Seq primer: T7 Promoter Primer. Location/Qualifiers
FEATURES	

```

source
1. .577
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCSPFL418OE09"
/lab_host="DH10B"
/clone_lib="FL4"
/note="Organ: Developed inflorescence and rachis (20cm-long); Vector: pSPORT; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [Developed inflorescence and rachis (20cm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucet.lad.ic.unicamp.br/public"

ORIGIN
Query Match 3.6%; Score 38.6; DB 14; Length 577;
Best Local Similarity 58.1%; Pred. No. 47;
Matches 68; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 587 CCGAGGAAGTTGGCGCAACTGGCAAGCGTAGCTTCCACGAGGCCACGACGACGCG 646
Db 537 CCTTGGAGGTTTGGCGGAACCGCGCGCGGTTCACACCCCGTAGCCGGCCACAG 478
QY 647 CCGAGCGCAACCCCGAGCGCGCTCCTCTGCAACAGCGGCTTCCGAACCTGAAGCGG 703
Db 477 CGCGCGTAGCTTCGCGCCCTCGCGCGGTGTAGCCACGCGCGGTAGCGCGCGCG 421

RESULT 14
CF877969
LOCUS
DEFINITION
tric077xp18.b1 T.reesei mycelial culture, Version 6 October 2003
ACCESSION
Hypocrea jecorina cDNA clone tric077xp18, mRNA sequence.
VERSION
CF877969.1 GI:38132651
KEYWORDS
EST.
SOURCE
Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
1 (bases 1 to 717)
Dienes S.E., Dankmeyer L., Dunn-Coleman N., Houfek T.D.,
Mitchell T.K., van Solingen P., Teunissen P.J.M., Ward M. and
Dean R.A.
REFERENCE
Analysis of the protein processing and secretion pathways in a
Trichoderma reesei EST dataset
Unpublished (2003)
COMMENT
Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph.dean@ncsu.edu
Seq primer: Lt-F1 primer.
Location/Qualifiers
1. .717
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/dev_stage="tric077xp18"
/clone_lib="T.reesei mycelial culture, Version 6 October 2003"
/note="Vector: pREP3Y; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
FEATURES
source

```

ORIGIN

Query Match 3.6%; Score 38.6; DB 14; Length 717;
Best Local Similarity 49.3%; Pred. No. 51;
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 585 GACCGAGGAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGGAGGCCACGACGACGAG 644
Db 488 GAAGAAGGCCAAGAAGGCGCAAGCGAAGGCGTTCAAGAGGCTCGTGCCCAACATGTCCCT 547

Qy 645 CGCGGAGGACACACCCGCGCGCTCACTGCAACCGCGCTCGGAACCGACCGACCGGGA 704
Db 548 CGGTGGGCAAGACCCAGGCTTTGACGCTCGGTCAACGCGCGCTCGGCGCGGTGT 607

Qy 705 ACATTTCACCTTCCCTGGCTAGAAATGGCGTGGATCATTTACGAACCGACACCGCCGAAA 764
Db 608 CCATTTCGCTGCTCGCGGCAACGACAGCGTGATGCTTGCAACTACTCCCGCGTGC 667

Qy 765 CGAAATTCAAAGCTTACTGTCGT 789
Db 668 CGCCACTGAGCCCTCACCGTCGT 692

RESULT 15
CF865608 718 bp mRNA linear EST 31-OCT-2003
LOCUS trico02xel17.b1 T.reesei mycelial culture, version 6 October 2003
DEFINITION Hypocrea jecorina cDNA clone trico02xel17, mRNA sequence.
ACCESSION CF865608
VERSION CF865608.1 GI:38120234
KEYWORDS EST.
SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE 1 (bases 1 to 718)
AUTHORS Diener S.E., Darkmeyer L., Dunn-Coleman N., Houfek T.D.,
Mitchell T.K., van Solingen P., Teunissen P.J.M., Ward M. and
Dean R.A.
TITLE Analysis of the protein processing and secretion pathways in a
Trichoderma reesei EST dataset
JOURNAL Unpublished (2003)
COMMENT Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph_dean@ncsu.edu
Seq primer: LT-F1 primer.
Location/Qualifiers
1..718
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="trico02xel17"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October 2003"
/note="Vector: pREP3Y; Site1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."

FEATURES
source
1..718
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="trico02xel17"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October 2003"
/note="Vector: pREP3Y; Site1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."

ORIGIN

Query Match 3.6%; Score 38.6; DB 14; Length 718;
Best Local Similarity 49.3%; Pred. No. 51;
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 585 GACCGAGGAAGTTGGCGCAACTGGCAAGCGTACGTTGCCACGGAGGCCACGACGACGAG 644
Db 254 GAAGAAGGCCAAGAAGGCGCAAGCGGCTTCAAGGCGCTCGTGCCCAACATGTCCCT 313

Search completed: March 8, 2004, 01:19:36
Job time : 3281 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 01:19:45 ; Search time 129 Seconds
(without alignments)
792.885 Million cell updates/sec

Title: US-09-993-777-7
Perfect score: 1863
Sequence: 1 MASLGLTALLAATLAPFGA.....VELVAIVNPSALSPDSIKM 362

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1863	100.0	362	2 AAW00635	AAW00635 ILTV glyc
2	1863	100.0	362	2 AAW06787	AAW06787 ILTV glyc
3	1863	100.0	362	5 ABG31711	ABG31711 Infectiou
4	156	8.4	384	2 AAR47236	AAR47236 Wild-type
5	152	8.2	355	2 AAW11475	AAW11475 Marek's d
6	152	8.2	355	2 AAW13105	AAW13105 Marek's d
7	148	7.9	364	2 AAW22999	AAW22999 Canine he
8	148	7.9	364	2 AAW72659	AAW72659 Canine he
9	148	7.9	364	4 AAB51316	AAB51316 Canine he
10	148	7.9	364	7 AAE39133	AAE39133 CHV PCG13
11	143.5	7.7	350	1 AAP70645	AAP70645 Pseudorab
12	143.5	7.7	350	2 AAR63143	AAR63143 Glycoprot
13	143.5	7.7	350	4 AAE05396	AAE05396 Pseudorab
14	143	7.7	317	2 AAR80637	AAR80637 Bovine he
15	142.5	7.6	350	4 AAU04970	AAU04970 Pseudorab
16	141.5	7.6	350	4 AAB82502	AAB82502 Pseudorab
17	140.5	7.5	380	2 AAR48063	AAR48063 Sequence
18	139	7.5	371	4 AAB78825	AB78825 Pseudorab
19	138	7.4	370	3 AAY32474	AAY32474 DNA encod
20	125	6.7	356	2 AAW03136	AAW03136 Herpesvir
21	125	6.7	356	2 AAW03323	AAW03323 Herpesvir
22	123.5	6.6	253	2 AAR22323	AAR22323 Marek Dis
23	119	6.4	423	7 ADC42841	ADC42841 REMAP pro
24	115	6.2	612	6 AAE33343	AAE33343 Mouse CD4
25	114	6.1	456	2 AAR33180	AAR33180 L3T4 muta

ALIGNMENTS

RESULT 1

AAW00635
ID AAW00635 standard; protein; 362 AA.

XX AAW00635;

XX AC AAW00635;

XX 16-OCT-2003 (revised)

DT 19-NOV-1996 (first entry)

XX ILTV glycoprotein gl.

XX Infectious laryngotracheitis virus; ILTV; herpesvirus; attenuation;

KW vector; vaccine; chicken; poultry; immunisation; glycoprotein gl.

XX Gallid herpesvirus 1.

OS W09508622-A1.

PN 30-MAR-1995.

XX 16-SEP-1994; 94WO-US010628.

XX 24-SEP-1993; 93US-00126597.

XX (SYTR) SYNTRO CORP.

XX Wild MA, Cochran MD;

XX WPI; 1995-139591/18.

XX N-PSDB; AAT33504.

XX Recombinant attenuated infectious laryngotracheitis virus - for use in

PT vaccines to protect poultry from infection from the virus, also methods

PT of distinguishing between vaccinated and naturally infected birds.

XX Example 1; Page 102-103; 177pp; English.

XX The gl gene, spanning nucleotides 9874-10962 of the unique short region

CC (AAT33504) of infectious laryngotracheitis virus (ILTV), codes for a

CC glycoprotein (AAW00635) of approx. 39,7535 mol.wt. The gl glycoprotein is

CC homologous to Varicella-Zoster gl. Deletion of the gl gene results in an

CC attenuated ILTV that is useful as a vaccine against ILTV disease in

CC chickens. Recombinant virus deleted for gl was safe in animal trials.

CC Deletion of the gl gene serves as a negative marker to distinguish

CC vaccines from infected animals. A gene coding for a foreign antigen may

CC be inserted into the gl gene to produce a recombinant multivalent

CC vaccine. (Updated on 16-OCT-2003 to standardise OS field)

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

SQ Sequence 362 AA;

Query Match 100.0%; Score 1863; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 1.5e-173;
 Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVVAPRPEATIQQLFFMPGQ 60
 DB 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVVAPRPEATIQQLFFMPGQ 60

QY 61 RPHKPYSGTVRVAFRSDITNQCQYQELSEERFENCSTRSSSVFVGCCKTETVTFASNRLTG 120
 DB 61 RPHKPYSGTVRVAFRSDITNQCQYQELSEERFENCSTRSSSVFVGCCKTETVTFASNRLTG 120

QY 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKPDIQVFAIQLSVYQFANTAATRGLYSKASC 180
 DB 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKPDIQVFAIQLSVYQFANTAATRGLYSKASC 180

QY 181 RPFGLPTVQLEAYLRTEESWRNQAAYVATEATTTSAEATTPPTVATSAEAEHFTFP 240
 DB 181 RPFGLPTVQLEAYLRTEESWRNQAAYVATEATTTSAEATTPPTVATSAEAEHFTFP 240

QY 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVAAVVSATIGLVIVISIVTRNMCPT 300
 DB 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVAAVVSATIGLVIVISIVTRNMCPT 300

QY 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPDSI 360
 DB 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPDSI 360

QY 361 KM 362
 DB 361 KM 362

RESULT 2
 AAW06787
 ID AAW06787 standard; protein; 362 AA.
 XX AC AAW06787;
 DT 16-OCT-2003 (revised)
 DT 02-JUN-1997 (first entry)
 XX ILTV glycoprotein g1.
 XX ILTV; vaccine; vector; attenuation; poultry;
 KW avian infectious bronchitis virus; Newcastle disease virus;
 KW infectious bursal disease virus of chickens; Marek's disease virus;
 KW herpesvirus; glycoprotein g1.
 XX Gallid herpesvirus 1; USDA strain 8302.
 XX FH Key Location/Qualifiers
 FT Peptide 1..22
 FT /label= Sig_peptide
 FT Protein 23..362
 FT /label= Mat_protein
 FT Region 272..292
 FT /label= Transmembrane_helix
 XX W09629396-A1.
 XX PD 26-SEP-1996.
 XX PF 21-MAR-1996; 96WO-US003916.
 XX PR 23-MAR-1995; 95US-00410121.
 XX PR 06-JUN-1995; 95US-00468190.
 XX PA (SYTR) SYNTRO CORP.
 XX PI Wild MA, Cochran MD;

XX WPI; 1996-443172/44.
 DR N-PSDB; AAT44384, AAT44385.
 XX Recombinant infectious laryngotracheitis virus with deletion in the glycoprotein G, g1 or US2 gene, etc. - useful for vaccines against infectious laryngotracheitis in poultry.
 PT
 PT
 XX Example 11; Page 110-111; 216pp; English.
 XX Glycoprotein g1 (AAW06787) is encoded by ORF8 of the unique short region (AAT44384) of infectious laryngotracheitis virus (ILTV). It shows homology to the varicella zoster virus g1 glycoprotein. Recombinant ILTV g1 protein produced in a swinepox virus reacts to convalescent sera from ILTV-infected chickens. Deletion of the g1 gene results in an attenuated ILTV that is useful as a vaccine and as a negative marker to distinguish CC vaccinated from infected animals. Insertion of a foreign gene into the g1 gene allows prodn. of multivalent vaccines. (Updated on 16-OCT-2003 to standardise OS field)
 CC
 CC
 XX Sequence 362 AA;
 Query Match 100.0%; Score 1863; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 1.5e-173;
 Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVVAPRPEATIQQLFFMPGQ 60
 DB 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVVAPRPEATIQQLFFMPGQ 60

QY 61 RPHKPYSGTVRVAFRSDITNQCQYQELSEERFENCSTRSSSVFVGCCKTETVTFASNRLTG 120
 DB 61 RPHKPYSGTVRVAFRSDITNQCQYQELSEERFENCSTRSSSVFVGCCKTETVTFASNRLTG 120

QY 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKPDIQVFAIQLSVYQFANTAATRGLYSKASC 180
 DB 121 PPHPKLTIRNPRNDGMYFVIVRLDDTKPDIQVFAIQLSVYQFANTAATRGLYSKASC 180

QY 181 RPFGLPTVQLEAYLRTEESWRNQAAYVATEATTTSAEATTPPTVATSAEAEHFTFP 240
 DB 181 RPFGLPTVQLEAYLRTEESWRNQAAYVATEATTTSAEATTPPTVATSAEAEHFTFP 240

QY 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVAAVVSATIGLVIVISIVTRNMCPT 300
 DB 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTVAAVVSATIGLVIVISIVTRNMCPT 300

QY 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPDSI 360
 DB 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPDSI 360

QY 361 KM 362
 DB 361 KM 362

RESULT 3
 ABG31711
 ID ABG31711 standard; protein; 362 AA.
 XX AC ABG31711;
 XX DT 29-AUG-2003 (revised)
 XX DT 15-NOV-2002 (first entry)
 XX Infectious laryngotracheitis virus (ILTV) glycoprotein I.
 XX Infectious laryngotracheitis virus; novel recombinant avian herpesvirus;
 KW NAHV; herpes virus of turkeys; HVT; Marek's disease; Newcastle disease;
 KW infectious laryngotracheitis; virucide; immunostimulant; vaccine; ILTV;
 KW glycoprotein I.
 XX Gallid herpesvirus 1.
 XX OS
 XX PI

PN US2002081316-A1.
 XX 27-JUN-2002.
 XX 14-JUN-2001; 2001US-00881457.
 XX 12-JUN-1992; 92US-00898087.
 XX 26-FEB-1993; 93US-00023610.
 XX 14-JUN-1993; 93WO-US0005681.
 XX 09-AUG-1994; 94US-00288065.
 XX 09-AUG-1995; 95WO-US010245.
 XX 13-JUN-1996; 96US-00663566.
 XX 21-FEB-1997; 97US-00804372.
 XX 23-OCT-1999; 99US-00426352.
 XX (COCH/) COCHRAN M D.
 XX (COOK/) COOK S M.
 XX (WILD/) WILD M A.
 XX Cochran MD, Cook SM, Wild MA;
 XX WPI; 2002-635456/68.
 XX N-PSDB; AEK90557.
 XX Novel recombinant avian herpesvirus comprising unique long and repeat
 PT viral genome regions of herpes virus of turkeys, unique short viral
 PT genome region of Marek's disease virus, and optional foreign DNA
 PT sequence.
 XX Disclosure; Page 21-22; 26pp; English.
 XX The invention relates to a novel recombinant avian herpesvirus (NAHV)
 CC comprising a herpes virus of turkeys (HVT) unique long (UL) and repeat
 CC viral genome region and a Marek's disease virus unique short (US) viral
 CC genome region where at least one foreign DNA sequence is inserted within
 CC a US2 gene of the US region of the NAHV, and is capable of being
 CC expressed in a host cell. NAHV is useful for producing vaccines used for
 CC immunising an avian species against Newcastle disease, infectious
 CC laryngotracheitis and Marek's disease. This sequence represents
 CC infectious laryngotracheitis virus (ILT) glycoprotein I. (Updated on 29-
 CC AUG-2003 to standardise OS field)
 XX SQ Sequence 362 AA;
 Query Match 100.0%; Score 1863; DB 5; Length 362;
 Best Local Similarity 100.0%; Pred. No. 1.5e-173;
 Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASLLGLTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPATLQLQFMPGQ 60
 DB 1 MASLLGLTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPATLQLQFMPGQ 60
 QY 61 RPHKPSGTVAFRSDITNOCYQELSEERFENCNTHSSSVFVGCCKVTEYTFASNRLTG 120
 DB 61 RPHKPSGTVAFRSDITNOCYQELSEERFENCNTHSSSVFVGCCKVTEYTFASNRLTG 120
 QY 121 PPHPKLTIRNPRNDGSMFVIVRLDDTKPEIDVFAIQLSVQFANTAATRGLYSKASC 180
 DB 121 PPHPKLTIRNPRNDGSMFVIVRLDDTKPEIDVFAIQLSVQFANTAATRGLYSKASC 180
 QY 181 RTFGLPTVQLAYLRTEESWRNQAAYATEATTTSAEATTPPTVATSAELEAEHFTFP 240
 DB 181 RTFGLPTVQLAYLRTEESWRNQAAYATEATTTSAEATTPPTVATSAELEAEHFTFP 240
 QY 241 WLENGVDHYETTPANENSVNTRVLTGMSPTLIGVTVAASVATIGLVISIVTRNMCPT 300
 DB 241 WLENGVDHYETTPANENSVNTRVLTGMSPTLIGVTVAASVATIGLVISIVTRNMCPT 300
 QY 301 HKLDTVSQDDEERSQTRRSKRGKPGMWACINKGADQDSSELVELVAIVNPSALSSPDSI 360
 DB 301 HKLDTVSQDDEERSQTRRSKRGKPGMWACINKGADQDSSELVELVAIVNPSALSSPDSI 360
 QY 361 KM 362

DB 361 KM 362
 RESULT 4
 ID AAR47236 standard; protein; 384 AA.
 XX AAR47236;
 XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 06-SEP-1994 (first entry)
 XX Wild-type Feline Herpes Virus ORF-1-encoded protein.
 XX Feline herpes virus; FHV; genome; mutant; vaccine; ORF-1;
 KW feline viral rhinotracheitis; herpesviridae family;
 KW herpes virus A subfamily.
 XX Feline herpesvirus.
 OS WO9403621-A1.
 XX 17-FEB-1994.
 PD 23-JUL-1993; 93WO-EP001971.
 PF 30-JUL-1992; 92EP-00202365.
 PR (ALKU) AKZO NV.
 PA Sondermeijer PJA, Willemse MJ;
 XX WPI: 1994-065709/08.
 DR N-PSDB; AAR47236.
 XX Feline herpes virus mutant comprising a heterologous gene inserted in the
 PT virus genome - for vaccination against feline herpes virus and other
 PT feline pathogens.
 XX Claim 2; Page 37-38; 55pp; English.
 XX Mutant versions of the wild-type feline herpes virus genome (AAR47236)
 CC are claimed. Esp. the FHV mutant is one which fails to produce one or
 CC more antigenic or functional polypeptides. The mutant may contain a
 CC heterologous nucleic acid sequence inserted within one of the 6 open
 CC reading frames. The attenuated FHV mutants can be used to prepare
 CC vaccines against feline viral rhinotracheitis and, where the heterologous
 CC insert encodes an antigen of a feline pathogen, against other infectious
 CC diseases of felines. (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX SQ Sequence 384 AA;
 Query Match 8.4%; Score 156; DB 2; Length 384;
 Best Local Similarity 23.3%; Pred. No. 2.9e-06;
 Matches 75; Conservative 48; Mismatches 141; Indels 58; Gaps 12;
 QY 5 LGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPATLQLQFMPGQ-RP 62
 DB 1 MSIAPIYILMAITGVYGVIRGDSVHSLVDTSFGVIYPTLENFTIYGHLLIFLDDQPLP 60
 QY 63 HRPYSGTVAVFRSDITNOCYQELSEERFENCNTHSSSVFVGCCKVTEYTFASNRLTG 120
 DB 61 VANNYNGTLEI-THYHSHSSYKIVQVIEYSSCFRVRNNAFRSCLHKTSHQDDQJSLNTS 119
 QY 121 PPHPKLTIRNPRNDGSMFVIVRLDDTKPEIDVFAIQLSVQFANTAATRGLYSKASC 180
 DB 120 VETGMLLTITSPKMGEGGIYALRVFRFNHNNK-ADVFGLSVFVYSP----DTGRHHRHAD- 173
 QY 181 RTFGLPTVQLAYLRTEESWRNQAAYATEATTTSAEATTPPTVATSAELEAEHFTFP 240

Db 174 -----ENLGEILTPS--SMETVVKVNTPIYDHMTTQT-----TSKMSSE-----P 216

Qy 241 WLENGVDHYEPTPANENSVTVRL-----GTMSPTLIGTVVAAVVSAT- 283

Db 217 SNTSISCHTFQNDPNEGELYTHLLNAGNIYYDDWMDGTTLQRLIDMGLNLSVTSSP 276

Qy 284 -----IGLVIVISI 292

Db 277 KNETTKQWTPDRKVGFIVISI 298

RESULT 5

AAW11475

ID AAW11475 standard; protein; 355 AA.

XX AAW11475;

XX 30-APR-1997 (first entry)

XX Marek's disease virus glycoprotein gI.

XX MDV; UL32; membrane glycoprotein 82; glycoprotein gI; antigen; vaccine;

XX vector; fowlpox virus.

XX Marek's disease gammaherpesvirus type 1 strain GA.

XX WO9703187-A2.

XX 30-JAN-1997.

XX OS-JUL-1996; 96WO-US011360.

XX 07-JUL-1995; 95US-00499474.

XX (JAPG) NIPPON ZEON KK.

XX (USDA) US SEC OF AGRIC.

XX Lee LP, Nazerian K, Witter RL, Wu P, Yanagida N, Yoshida S;

XX WPI; 1997-119044/11.

XX New DNA encoding glyco:protein 82 of Marek disease virus - useful in

XX vaccines to protect poultry.

XX Disclosure; Page 72-73; 101pp; English.

XX Genes encoding Marek's disease virus glycoprotein gI (AAW11475) and/or

XX glycoprotein gE (AAW11474) can be incorporated into recombinant viral

XX vectors that also carry the UL32 gene (see also AAT51358) encoding

XX glycoprotein 82 (AAW11473). The gI and gE genes were cloned from MDV

XX genomic DNA by PCR (see also AAT51368-69 and AAT51366-67). A transfer

XX vector was constructed that was used to produce a recombinant fowlpox

XX virus useful as a vaccine to protect poultry against MDV infection

XX

SQ Sequence 355 AA;

Query Match 8.2%; Score 152; DB 2; Length 355;

Best Local Similarity 22.9%; Pred. No. 6.5e-06;

Matches 78; Conservative 55; Mismatches 166; Indels 42; Gaps 14;

Qy 19 GAMGIVITGNHVSARIIDDDHIVIVAPRPE-ATIQQLFFMPGQRPKPKYSGTVRVAFRSD 77

Db 15 GIWSIVTGTSTVLSTQDQALVAFGLDKMNVNRGQLFLGDTQRTSSYTGTTET-LKWD 73

Qy 78 ITNOCYQELSERFENCNTHRSSSVFVGCKVTEYTFASNLGTGPPHFK---LTIRNRP 134

Db 74 BEYKCYSLVHATSYMDCPAIDATVFRGCRDAVVAQAQPHDRV--QPFPEKGTLLRIVEPRV 131

Qy 135 NDSGMFYIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASCTFGLPTVQLEAYL 194

Db 132 SDTGSYIRVALAG-RNMSDIFRMAVIRS-----SKSWACHNSASSFOAHKCI 179

Qy 195 RTEE--SWRNQAYVATEATTTSABATTPVTATTSASELEAEHFTFPWLENGVDHYEPT 252

Db 180 RYDRAFAFENYLIGHVGNLLDSSELHAYNITPQIS-TDINIITFFYDNGSTIYSPT 238

Qy 253 P---ANENSNTVRLGT-MSPTLIGTVVAAV--SATIGLVIV---ISIVTRNMCTPHRK 303

Db 239 VENLFNNSHVDAMNSTGMNTVLKVTLPRLIYFSTMIVLCTIALAIVLYVCERCSPHER 298

Qy 304 LDTVSQDDERSQTRRESRKFGPMVACEINKGADQDSELVE 344

Db 299 I---YICEPRSD-----APLITSAVNESFOYDYNVKE 328

RESULT 6

AAW13105

ID AAW13105 standard; protein; 355 AA.

XX AAW13105;

XX 17-OCT-2003 (revised)

XX 12-MAY-1997 (first entry)

XX Marek's disease virus type 1 glycoprotein I.

XX GA strain; glycoprotein; gPI; gpE; recombinant virus;

XX Newcastle disease virus; herpes virus; vaccine.

XX Marek disease virus type 1.

XX JP09009978-A.

XX 14-JAN-1997.

XX 26-APR-1996; 96JP-00131084.

XX 28-APR-1995; 95JP-00129523.

XX (JAPG) JAPANESE GEON CO LTD.

XX WPI; 1997-126432/12.

XX N-PSDB; AAT61883.

XX Recombinant virus contg. herpes virus glyco:protein gE - and effective

XX Newcastle disease vaccine contg. it.

XX Claim 5; Page 14-15; 20pp; Japanese.

XX The present sequence encodes the Marek's disease virus (MDV) type 1 (GA

XX strain) glycoprotein (gp) I. A recombinant virus containing the gene

XX encoding MDV gpE, and optionally also gpI, in a genomic region

XX unessential for its growth, can be used as the effective component in a

XX Newcastle disease virus vaccine. (Updated on 17-OCT-2003 to standardise

XX OS field)

XX Sequence 355 AA;

Query Match 8.2%; Score 152; DB 2; Length 355;

Best Local Similarity 22.9%; Pred. No. 6.5e-06;

Matches 78; Conservative 55; Mismatches 166; Indels 42; Gaps 14;

Qy 19 GAMGIVITGNHVSARIIDDDHIVIVAPRPE-ATIQQLFFMPGQRPKPKYSGTVRVAFRSD 77

Db 15 GIWSIVTGTSTVLSTQDQALVAFGLDKMNVNRGQLFLGDTQRTSSYTGTTET-LKWD 73

Qy 78 ITNOCYQELSERFENCNTHRSSSVFVGCKVTEYTFASNLGTGPPHFK---LTIRNRP 134

Db 74 BEYKCYSLVHATSYMDCPAIDATVFRGCRDAVVAQAQPHDRV--QPFPEKGTLLRIVEPRV 131

Qy 135 NDSGMFYIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASCTFGLPTVQLEAYL 194

Db 132 SDTGSYIRVALAG-RNMSDIFRMAVIRS-----SKSWACHNSASSFOAHKCI 179

Qy 195 RTEE--SWRNQAYVATEATTTSABATTPVTATTSASELEAEHFTFPWLENGVDHYEPT 252

```
Db 180 RYVDRMAFENYLIGHVGNLSDSELHAIYNIPTQGIS-TDINIITPPYDMSGTIYSPT 238
QY 253 P----ANENSNTVRLGN-MSPTLIGVTAAV- -SATIGLVIV- -ISIVTRNMCPTHEK 303
Db 239 VFNLFNNSHVDAMNSTMGMNTVLKYLPLRIYFTSWIVLCIITAIYLVCEKCRSPHRR 298
QY 304 LDTVSQDDEBSQTRRESKFGPMVACEINKGADQDSSELVE 344
Db 299 I- -YIGERSDE- - - - -APLITSAVNESFYDYNVKE 328

RESULT 7
AAW22999
ID AAW22999 standard; protein; 364 AA.
XX
AC AAW22999;
XX
XX
DT 20-FEB-1998 (first entry)
XX
XX Canine herpesvirus glycoprotein I PCgI364.
XX
KW Vaccine; vector; gene therapy; canid; dog; CHV; CgI; PCgI364;
KW glycoprotein I.
XX
OS Canine herpesvirus.
XX
XX WO929772-A1.
XX
XX 21-AUG-1997.
XX
XX 14-FEB-1997; 97WO-US004115.
XX
XX 15-FEB-1996; 96US-00602010.
XX
XX (HESK-) HESKA CORP.
XX
XX Haanes EU, Frank RA;
XX
XX WPI; 1997-424758/39.
XX
XX N-PSDB; AAT75616.
XX
XX Recombinant canine herpes virus and its genome - useful as vaccine to
XX protect canids against infectious, metabolic or genetic diseases.
XX
XX Claim 57; Page 181-182; 240pp; English.
XX
XX This protein comprises canine herpesvirus (CHV) glycoprotein I (PCgI364).
XX Its sequence was deduced from a coding region found in CHV genomic DNA
XX molecule NCUS10592 (see AAT75616). PCgI364 can be expressed in
XX transformed host cells. CHV proteins, nucleic acids, and antibodies
XX raised against CHV proteins, can be used to protect canids against CHV
XX infection. Novel recombinant CHV and novel recombinant CHV genomes that
XX contain heterologous nucleic acid molecules inserted e.g. into CgI gene,
XX can be used as vaccines to protect canids against infectious, metabolic
XX or genetic diseases
XX
XX Sequence 364 AA;
XX
XX Query Match 7.9%; Score 148; DB 2; Length 364;
XX Best Local Similarity 27.3%; Pred. No. 1.7e-05;
XX Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;
XX
QY 4 LLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQLQFFMPQQR-P 62
Db 16 LLITMFLPILFLFYGVNGFYKGTIYISMFLNTSSGFSIFPDDKFIVSGRLFLDDQHL 75
QY 63 HKPSGTVRVAFRSDITNQCYQELSEERFENCNTHRSSSVFVGC--KVTEYTFASNRLTG 120
Db 76 VNNYSGTIEFTI--HFNNSCVTYQTIEYFSCPIFNNAFRSCLKVKSKHHESQLRINS 132
QY 121 PPHPFKLIRNPRNDSGMFVIVRLDDTKEPIDVFAIQLSVYQF 165
Db 133 IENGVLLEITNPKNDSGVFIQVLENNK--TDFGIPAFIYSF 175

RESULT 9
AAW22999
ID AAW22999 standard; protein; 364 AA.
XX
AC AAW22999;
XX
XX
DT 06-AUG-2003 (revised)
XX
```

```
RESULT 8
AAW72659
ID AAW72659 standard; protein; 364 AA.
XX
AC AAW72659;
XX
XX
DT 27-AUG-2003 (revised)
DT 07-JAN-1999 (first entry)
XX
XX Canine herpes virus protein sequence PCgI-364.
XX
XX Canine herpes virus; CHV; recombinant canine herpes virus vector; genome;
XX vaccination; dog; protozoan; helminth; ectoparasite; bacteria;
XX virus infection.
XX
XX Canine herpesvirus.
XX
XX US5804197-A.
XX
XX 08-SEP-1998.
XX
XX 12-JUL-1996; 96US-00680726.
XX
XX 15-FEB-1996; 96US-00602010.
XX
XX (HESK-) HESKA CORP.
XX
XX Frank RS, Haanes EU;
XX
XX WPI; 1998-505590/43.
XX
XX N-PSDB; AAV66941.
XX
XX Canine herpes virus nucleic acids - useful for producing recombinant
XX canine herpes virus vectors.
XX
XX Claim 7; Col 133-136; 103pp; English.
XX
XX The present sequence represents a canine herpes virus (CHV) protein
XX sequence. The nucleic acid sequence which encodes the protein can be used
XX for DNA vaccination of dogs against CHV and also other infective agents
XX such as protozoans, helminths, ectoparasites, bacteria and viruses. CHV
XX can be formulated by incorporation of heterologous nucleic acid molecules
XX as a single multivalent therapeutic composition against a variety of
XX canine pathogens. (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 364 AA;
XX
XX Query Match 7.9%; Score 148; DB 2; Length 364;
XX Best Local Similarity 27.3%; Pred. No. 1.7e-05;
XX Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;
XX
QY 4 LLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQLQFFMPQQR-P 62
Db 16 LLITMFLPILFLFYGVNGFYKGTIYISMFLNTSSGFSIFPDDKFIVSGRLFLDDQHL 75
QY 63 HKPSGTVRVAFRSDITNQCYQELSEERFENCNTHRSSSVFVGC--KVTEYTFASNRLTG 120
Db 76 VNNYSGTIEFTI--HFNNSCVTYQTIEYFSCPIFNNAFRSCLKVKSKHHESQLRINS 132
QY 121 PPHPFKLIRNPRNDSGMFVIVRLDDTKEPIDVFAIQLSVYQF 165
Db 133 IENGVLLEITNPKNDSGVFIQVLENNK--TDFGIPAFIYSF 175

RESULT 9
AAW22999
ID AAW22999 standard; protein; 364 AA.
XX
AC AAW22999;
XX
XX
DT 06-AUG-2003 (revised)
XX
```



```

XX AAP70645;
AC
XX 29-APR-1991 (first entry)
DT
XX
XX Pseudorabies virus gp63 protein.
DE
XX PRV; vaccine; gI; gp50; gp63.
KW
XX Pseudorabies virus.
OS
XX WO8702058-A.
PN
XX 09-APR-1987.
PD
XX
XX 28-AUG-1986; 86WO-US001761.
PF
XX
XX 04-OCT-1985; 85US-00784787.
PR
XX 26-NOV-1985; 85US-00801799.
PR
XX 26-MAR-1986; 86US-00844113.
PR
XX 26-MAR-1986; 86US-00844133.
PR
XX 18-JUL-1986; 86US-00886260.
PR
XX
XX (UPJO ) UPJOHN CO.
PA
XX (PETR/) PETROVSKIS E A.
PA
XX
XX Post LE, Petrovskis EA, Timmins JG;
PI
XX
XX WPI; 1987-108699/15.
DR
XX N-PSDB; AAN70994.
DR
XX
XX Pseudo-rabies virus protein - produced from recombinant DNA and used to
PT produce vaccine and detect animals infected with virulent virus.
PT
XX
XX Claim 8; Page 59; 67pp; English.
XX
XX The PRV glycoprotein product may be used as in vaccination of animals
CC such as swine, sheep and goats against infection by the virus. The
CC protein may be produced from a transformed expression system such as
CC E.coli, yeast or CHO cells operatively linked to a suitable expression
CC control sequence. Glycoproteins gpi and gp63 may be used to distinguish
CC between infected and vaccinated animals. See also AAN70993-5
CC
XX
XX Sequence 350 AA;
SQ
Query Match 7.7%; Score 143.5; DB 1; Length 350;
Best Local Similarity 22.8%; Pred. No. 4.3e-05;
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;
QY 8 LALLAATLAPFGAM--GIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQRPHPK 65
DB 15 LLLAALTTLAALTFRVGVLFPGAGSVHVAGSAVLVPGDAPNLTDGTLTLLFEGSPSN- 73
QY 66 YSGTVRVAFRSDITNQCYQLSEERFENCNTHRSSSVFVGC--KVTEYTFSSA-----NRL 118
DB 74 YSGRVEL-LRLDPKRCACYTREYAAEYDLCPRVHHEAFRGCLRKREPLARRASAAVEARRL 132
QY 119 TGPHPFKLIRPRNDSCMFYIVRLDDTKPEIDVFAIQLSVYQFANTATRGL---- 174
DB 133 -----LFVSRPAPPDAGSYVLRVNGT---TDLFVLTLV-----PPRGRPHHP 174
QY 175 --YSKASCRFTGLPTVQLEAYLRTEESWRNQAQV--ATEATTSA-----EATPTPTV 224
DB 175 TPSSADECR----FVV-----GSHDSLVRVDPDAEDAVTTTPTTPTTAPP 221
QY 225 TATSAS-----ELEAEHTFPWLENGVDHYEFTPANENSVTVRLGWNSTPLIGTVTA 277
DB 222 RGTGATPEPRSDDEEDE-----EGATTAMPVPGTLDANGTM---VLNASVWSRVLL 271
QY 278 AVVSATIG-----LVIVISIVRNWCTPHRKLDTVSDDEERSQTRRESK 323
DB 272 AAANATAGAPGPKIAMVLGPTIIVLLIFLGGVACAAARCGIASTGRDPGAARSTRR 331

```

RESULT 12
 AAR63143
 ID AAR63143 standard; protein; 350 AA.
 XX
 AC AAR63143;
 XX
 XX 25-MAR-2003 (revised)
 DT 12-JUN-1995 (first entry)
 XX
 XX Glycoprotein 63 (gp63) of pseudorabies virus.
 DE
 XX pseudorabies virus; PRV; glycoprotein; gp50; gI; gp63; determination;
 KW detection; vaccine; infected animal; isolation; cloning; virulence.
 KW
 XX Pseudorabies virus.
 OS
 XX US5352575-A.
 PN
 XX 04-OCT-1994.
 PD
 XX 20-APR-1990; 90US-00513282.
 PF
 XX 04-OCT-1985; 85US-00784787.
 PR
 XX 26-NOV-1985; 85US-00801799.
 PR
 XX 26-MAR-1986; 86US-00844113.
 PR
 XX 16-JUL-1986; 86US-00886260.
 PR
 XX 29-JUN-1987; 87US-00100817.
 PR
 XX (UPJO) UPJOHN CO.
 PA
 XX Timmins JG, Post LE, Petrovskis EA;
 PI
 XX WPI; 1994-316176/39.
 DR
 XX N-PSDB; AAR73489.
 DR
 XX
 XX Identifying animals vaccinated against pseudorabies virus - by detecting
 PT the absence of gI or gp63 antibodies in serum to distinguish vaccinated
 PT from infected animals.
 PT
 XX Example 3; Col 21-24; 21pp; English.
 PS
 XX AAR63143 shows the protein sequence of gp63 (glycoprotein) encoded by
 CC AAR73489, isolated from Pseudorabies virus (PRV). The DNA and protein
 CC sequences of the invention are useful in a method for distinguishing an
 CC animal vaccinated with a PRV vaccine lacking glycoprotein gI, gp63 or
 CC gp50 from an animal infected with a virulent wild-type PRV without
 CC sacrificing the animal. The method is used to test animals such as swine,
 CC cattle, sheep and goats. (see AAR63142 and AAR63144). (Updated on 25-MAR-
 CC 2003 to correct PF field.)
 CC
 XX Sequence 350 AA;
 SQ

Query Match 7.7%; Score 143.5; DB 2; Length 350;
 Best Local Similarity 22.8%; Pred. No. 4.3e-05;
 Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;
 QY 8 LALLAATLAPFGAM--GIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQRPHPK 65
 DB 15 LLLAALTTLAALTFRVGVLFPGAGSVHVAGSAVLVPGDAPNLTDGTLTLLFEGSPSN- 73
 QY 66 YSGTVRVAFRSDITNQCYQLSEERFENCNTHRSSSVFVGC--KVTEYTFSSA-----NRL 118
 DB 74 YSGRVEL-LRLDPKRCACYTREYAAEYDLCPRVHHEAFRGCLRKREPLARRASAAVEARRL 132
 QY 119 TGPHPFKLIRPRNDSCMFYIVRLDDTKPEIDVFAIQLSVYQFANTATRGL---- 174
 DB 133 -----LFVSRPAPPDAGSYVLRVNGT---TDLFVLTLV-----PPRGRPHHP 174
 QY 175 --YSKASCRFTGLPTVQLEAYLRTEESWRNQAQV--ATEATTSA-----EATPTPTV 224
 DB 175 TPSSADECR----FVV-----GSHDSLVRVDPDAEDAVTTTPTTPTTAPP 221

QY 225 TATSAS-----PLEAEHFTFPWLENGVDHYETPANENSVTVRLGTMSTPLIGTVVA 277
Db 222 RGTGATPEPRSDDEEED-----EGATTAMTPVPGTLDANGTM---VLNASVSVRVL 271
QY 278 AVVSATIG-----LVIVISIVTRNMCTPHRKLDTVSDDEERSQTRRESRK 323
Db 272 AAANATAGARGPGKIAMVLPVLLIFLGVCACARRCARGIASTGRDPGARRSTR 331

RESULT 13
AAE05396
ID AAE05396 standard; protein; 350 AA.
AC AAE05396;
XX
DT 24-SEP-2001 (first entry)
XX
XX Pseudorabies virus (PRV) glycoprotein gp63.
DE Pseudorabies virus; PRV; glycoprotein; gp63; immunostimulant; vaccine;
KW PRV infection.
XX
XX Pseudorabies virus.
XX
XX Key Location/Qualifiers
FH Misc-difference 131 /note= "Encoded by G3G"
FT Misc-difference 132 /note= "Encoded by GTG"
FT Misc-difference 144 /note= "Encoded by GCG"
FT Misc-difference 207 /note= "Encoded by GCG"
FT Misc-difference 222 /note= "Encoded by GCG"
FT Misc-difference 223 /note= "Encoded by GCG"
FT Misc-difference 284 /note= "Encoded by GCG"
FT Misc-difference 334 /note= "Encoded by GCG"
FT /note= "Encoded by GCC"
XX
XX US6261563-B1.
XX
PD 17-JUL-2001.
XX
XX 07-JUN-1995; 95US-00485287.
XX
PR 26-NOV-1985; 85US-00801799.
PR 26-MAR-1986; 86US-00844113.
PR 16-JUL-1986; 86US-00886260.
PR 28-AUG-1986; 86WO-US001761.
PR 29-JUN-1987; 87US-00100817.
PR 20-APR-1990; 90US-00513282.
PR 21-JUN-1994; 94US-00262813.
XX
XX (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Petrovskis EA, Post LR, Timmins JG;
XX
DR WPI; 2001-450478/48.
DR N-PSDB; AAD10195.
XX
XX Polypeptide comprising a pure and isolated pseudorabies virus gp63
PT polypeptide or its fragments, useful as a vaccine for protecting animals
PT against pseudorabies virus infection.
XX
XX Claim 1; Col 39-40; 21pp; English.
XX
XX The invention relates to Pseudorabies virus (PRV) glycoproteins gp50,
CC gp63, gI and their corresponding DNA molecules. These glycoproteins are
CC used as vaccines for protecting animals against PRV infection. The
CC invention also relates to methods for protecting animals against PRV

CC infections and methods for distinguishing between infected and vaccinated
CC animals. The present sequence is Pseudorabies virus (PRV) glycoprotein
CC gp63
XX
SQ Sequence 350 AA;
Query Match 7.7%; Score 143.5; DB 4; Length 350;
Best Local Similarity 22.8%; Fred. No. 4.3e-05;
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;
QY 8 LALLAATLAPFCAM--GIVITGNHVSARIDDDHIVIVAPRPEATIQLOLFMPGQRPKHP 65
Db 15 LLLAALTTLAALTFRVGGVLPFGAGSVHVAGSAVLVPGDAPNLITDGLLFLGSPSN- 73
QY 66 YSGTVRVAFRSDITNQCYQELSEERFENCSTRSSSVFVGC--KVTEYTFSSA-----NRL 118
Db 74 YSRVEL-LRLDPKEACYTREVAAYDLCPRVHEAPRGCLRKREPLARRASAAVEARRL 132
QY 119 TQPPHFKLTRNRPNDSGMFYVIVLDDTKBIDVFAIQLSVYQFANTATRGL----- 174
Db 133 -----LFVSRPAPPDAGSVVLRVRVNGT---TDLFVLTVLV-----PPRGRPHHP 174
QY 175 --YSKASCTFGLPTVQLEAYLRTESWRNQAVV--ATEATTTS-----AEATTPTPV 224
Db 175 TPSSADECR---PVV-----GSHDSLVRVVDPAEDAVFTTPPIEPPTTPAPP 221
QY 225 TATSAS-----ELEAEHFTFPWLENGVDHYETPANENSVTVRLGTMSTPLIGTVVA 277
Db 222 RGTGATPEPRSDDEEED-----EGATTAMTPVPGTLDANGTM---VLNASVSVRVL 271
QY 278 AVVSATIG-----LVIVISIVTRNMCTPHRKLDTVSDDEERSQTRRESRK 323
Db 272 AAANATAGARGPGKIAMVLPVLLIFLGVCACARRCARGIASTGRDPGARRSTR 331

RESULT 14
AAR80637
ID AAR80637 standard; protein; 317 AA.
XX
AC AAR80637;
XX
DT 21-DEC-1995 (first entry)
XX
DE Bovine herpes virus-1 mutant ORF1 product.
XX
XX BHV-1; vaccine; gIV gene; antigen.
XX
XX Bovine herpesvirus.
XX
XX EP663403-A1.
XX
XX 19-JUL-1995.
XX
XX 18-NOV-1994; 94EP-00203361.
XX
XX 23-NOV-1993; 93EP-00203274.
XX
XX (ALKU) AKZO NOBEL NV.
XX
XX Keil G;
XX
XX WPI; 1995-247499/33.
DR N-PSDB; AAQ99001.
XX
XX New bovine herpes virus with mutation in ORF1 and polyA signal for gIV
PT gene - useful as a vaccine, esp. where the mutation is insertion of
PT sequence encoding an heterologous antigen.
XX
XX Claim 2; Page 13-14; 28pp; English.
XX
XX To produce a new BHV-1 mutant, DNA encoding gIV (flanked on both sides by
CC a sequence that allows homologous recombination) and a genotypically-neg.
CC BHV-1 (or its DNA) is introduced into a host cell. Since the gIV-neg.

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 02:39:45 ; Search time 30 Seconds
(without alignments)
622.953 Million cell updates/sec

Title: US-09-993-777-7
Perfect score: 1863
Sequence: 1 MASLIGTALLAATLAPFGA.....VELVAIVNPSALSPDSIKM 362

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1863	100.0	362	5	PCT-US96-03916-7
2	1863	100.0	362	5	PCT-US96-03916-68
3	156	8.4	384	4	US-08-504-617-2
4	152	8.2	355	4	US-08-709-731A-28
5	148	7.9	317	2	US-08-344-833-2
6	148	7.9	364	1	US-08-680-726A-56
7	148	7.9	364	3	US-09-092-409-56
8	145	7.8	370	3	US-08-911-321-10
9	143.5	7.7	350	6	5352575-7
10	140.5	7.5	380	3	US-08-924-345-2
11	112	6.0	387	4	US-09-175-928-2
12	108	5.8	713	1	US-08-190-802A-63
13	108	5.8	713	3	US-08-477-346-63
14	108	5.8	713	4	US-08-473-089-63
15	108	5.8	713	4	US-08-487-072A-63
16	108	5.8	713	4	US-09-108-857-3
17	107.5	5.8	100	3	US-09-047-125-27
18	107.5	5.8	100	2	US-07-736-335B-27
19	103.5	5.6	94	1	US-08-150-203A-6
20	103.5	5.6	94	1	US-08-454-730-6
21	103.5	5.6	94	4	US-08-949-788-6
22	99	5.3	97	2	US-09-047-125-25
23	99	5.3	97	3	US-07-736-335B-25
24	99	5.3	820	4	US-09-489-039A-10142
25	98	5.3	503	1	US-07-946-497-2
26	98	5.3	503	1	US-08-483-322-2
27	98	5.3	503	2	US-08-478-882-2

28 97 5.2 1220 2 US-08-843-530B-36 Sequence 36, Appl
29 94.5 5.1 307 4 US-09-197-970B-3 Sequence 3, Appl
30 93.5 5.0 878 4 US-09-556-706B-2 Sequence 2, Appl
31 93.5 5.0 907 3 US-08-783-774-2 Sequence 2, Appl
32 93.5 5.0 907 4 US-09-328-599A-1 Sequence 1, Appl
33 93.5 5.0 907 5 PCT-US95-04611A-19 Sequence 19, Appl
34 92.5 5.0 547 1 US-08-314-615-1 Sequence 1, Appl
35 92.5 5.0 547 1 US-08-314-362-1 Sequence 1, Appl
36 92.5 5.0 547 1 US-08-433-010-1 Sequence 1, Appl
37 92.5 5.0 547 1 US-08-482-882-1 Sequence 1, Appl
38 92.5 5.0 547 2 US-08-483-389-1 Sequence 1, Appl
39 92.5 5.0 547 2 US-08-487-113D-1 Sequence 1, Appl
40 92.5 5.0 547 2 US-08-473-503-1 Sequence 1, Appl
41 92.5 5.0 547 2 US-08-483-932-1 Sequence 1, Appl
42 92.5 5.0 547 2 US-08-720-420A-1 Sequence 1, Appl
43 92.5 5.0 547 3 US-08-714-017-1 Sequence 1, Appl
44 92.5 5.0 547 3 US-08-863-790-1 Sequence 1, Appl
45 92.5 5.0 547 3 US-08-475-680-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
PCT-US96-03916-7
; Sequence 7, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-03916-7

Query Match 100.0%; Score 1863; DB 5; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.3e-188;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLIGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVAVRPPEATIQQLFFMFGQ 60
DB 1 MASLIGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVAVRPPEATIQQLFFMFGQ 60

QY 61 RPKPXSQTVRVAFRSDITNQCYQELSEERFENCNTHRSSVFGVCKVTEYTSASNRLTG 120
DB 61 RPKPXSQTVRVAFRSDITNQCYQELSEERFENCNTHRSSVFGVCKVTEYTSASNRLTG 120
QY 121 PPFPFKLTIRNPRNDGSMFYVIRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
DB 121 PPFPFKLTIRNPRNDGSMFYVIRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
QY 181 RTFGLPTVQLVQLAYLRTESWRNWOAYVATEATTTSABATTPPVVATTSASELEAEHFTFP 240
DB 181 RTFGLPTVQLVQLAYLRTESWRNWOAYVATEATTTSABATTPPVVATTSASELEAEHFTFP 240
QY 241 WLENGVDHYEPTPANENSNVTVRGTMSPTLIGVTVAAVSATIGLVIVISIVTRNMCCTP 300
DB 241 WLENGVDHYEPTPANENSNVTVRGTMSPTLIGVTVAAVSATIGLVIVISIVTRNMCCTP 300
QY 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPPDSI 360
DB 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPPDSI 360
QY 361 KM 362
DB 361 KM 362

RESULT 2

PCT-US96-03916-68
Sequence 68, Application PC/TUS9603916
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-03916-68

Query Match 100.0%; Score 1863; DB 5; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.3e-188;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHHIVIVAPRPEATIQQLFFPMGQ 60
DB 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHHIVIVAPRPEATIQQLFFPMGQ 60
QY 61 RPKPXSQTVRVAFRSDITNQCYQELSEERFENCNTHRSSVFGVCKVTEYTSASNRLTG 120
DB 61 RPKPXSQTVRVAFRSDITNQCYQELSEERFENCNTHRSSVFGVCKVTEYTSASNRLTG 120
QY 121 PPFPFKLTIRNPRNDGSMFYVIRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
DB 121 PPFPFKLTIRNPRNDGSMFYVIRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
QY 181 RTFGLPTVQLVQLAYLRTESWRNWOAYVATEATTTSABATTPPVVATTSASELEAEHFTFP 240
DB 181 RTFGLPTVQLVQLAYLRTESWRNWOAYVATEATTTSABATTPPVVATTSASELEAEHFTFP 240
QY 241 WLENGVDHYEPTPANENSNVTVRGTMSPTLIGVTVAAVSATIGLVIVISIVTRNMCCTP 300
DB 241 WLENGVDHYEPTPANENSNVTVRGTMSPTLIGVTVAAVSATIGLVIVISIVTRNMCCTP 300
QY 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPPDSI 360
DB 301 HRKLDTVSQDDEERSQTRRESKFGPMVACEINKGADQDSSELVELVAIVNPSALSPPDSI 360
QY 361 KM 362
DB 361 KM 362

RESULT 3

US-08-504-617-2
Sequence 2, Application US/08504617
Patent No. 6521236
GENERAL INFORMATION:
APPLICANT: Willemse, Martha Jacoba
APPLICANT: Sondermeijer, Paulus Jacobus Antonius
TITLE OF INVENTION: Vector vaccines of recombinant
TITLE OF INVENTION: Feline herpesviruses
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 6521236el
STREET: 1330 Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/504,617
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/211,150
FILING DATE: March 22, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
OTHER INFORMATION: /label= ORF-1
US-08-504-617-2

TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-344-833-2

Query Match 7.9%; Score 148; DB 2; Length 317;
Best Local Similarity 27.1%; Pred. No. 4.1e-07;
Matches 68; Conservative 34; Mismatches 125; Indels 24; Gaps 11;

QY 1 MASLLGTLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLQFFMP 58
DB 1 MERLLLMVLAARAAP--ARSLVYRGEAVGLRADGPVAFVAFH-TDATALRGRLLFLE 57
QY 59 GQRP-HKPSGTIVRVAFRSDITNQCQELSEERFENCTHRSSSVFVGC--KVTEYTFASNLTG 117
DB 58 HQLPAGRRYNGTVEL-LRYHAAGDCFVMLQTTAFASCPRVANDAFRSLHADTRPARSER 116
QY 118 LTGP--PHPFKLTIRPNRNDSGMFYIVRL---DDTKPIDVFAIQLSVYQFAN---- 167
DB 117 RASAIVENHVLFSIARPIIDSLGLYFLRVGYGTAGSERRRDVFPPLAAVHSGFGGDP 176
QY 168 --TAATRGYS--KASCTFGLTIVQLEAYLRTESWRNWOAYVATEATTTSAAETPTP 223
DB 177 EAAAHAPGTVAEVACRER-GLDASSASLYDRALAAPGAA--TTPGPTASSESQAATP 233
QY 224 VTATSASELEA 234
DB 234 ERVDTEVEA 244

RESULT 6

US-08-680-726A-56
; Sequence 56, Application US/08680726A
; Patent No. 5804197
; GENERAL INFORMATION:
; APPLICANT: Haanes, Elizabeth J.
; APPLICANT: Frank, Rexann S.
; TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 12-JUL-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-46-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-680-726A-56

Query Match 7.9%; Score 148; DB 1; Length 364;
Best Local Similarity 27.3%; Pred. No. 5.1e-07;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLQFFMPGQR-P 62
DB 16 LLITMFLPILFLYGVNGFVYKGTIYSMTINTSSGFSIFPDDKFIIVSGRLLFLDDQHLS 75
QY 63 HKPSGTIVRVAFRSDITNQCQELSEERFENCTHRSSSVFVGC--KVTEYTFASNLTG 120
DB 76 VNNYSGTIEFI---HFNNSCYTVYQTIETFPFNNAPRSLCKKVKSHESQLRINS 132
QY 121 PPHFPLKLTIRPNRNDSGMFYIVRLDDTKPIDVFAIQLSVYQF 165
DB 133 LENGVLLEITNPKNDSGVFIIVQLENNK--TDVFGIPAFIYSP 175

RESULT 7

US-09-092-409-56
; Sequence 56, Application US/09092409
; Patent No. 6159478
; GENERAL INFORMATION:
; APPLICANT: Haanes, Elizabeth J.
; APPLICANT: Frank, Rexann S.
; TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 09/09/92, 409
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/680,726
; FILING DATE: 12-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-46-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-092-409-56

Query Match 7.9%; Score 148; DB 3; Length 364;
Best Local Similarity 27.3%; Pred. No. 5.1e-07;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLQFFMPGQR-P 62
DB 16 LLITMFLPILFLYGVNGFVYKGTIYSMTINTSSGFSIFPDDKFIIVSGRLLFLDDQHLS 75
QY 63 HKPSGTIVRVAFRSDITNQCQELSEERFENCTHRSSSVFVGC--KVTEYTFASNLTG 120
DB 76 VNNYSGTIEFI---HFNNSCYTVYQTIETFPFNNAPRSLCKKVKSHESQLRINS 132

QY 121 PPHPFKLTIRNPRNDGMPYVIVRLDDTKBPIDVFAIQLSVQF 165
DB 133 IENGVLLEITNPXPNDSGVFIRVQLENNK--TDVFGIPAFIYSF 175

RESULT 8
US-08-911-321-10
Sequence 10, Application US/08911321
Patent No. 6010703
GENERAL INFORMATION:
APPLICANT: Roger K. Maes and Stephen J. Spatz
TITLE OF INVENTION: Recombinant Poxvirus
TITLE OF INVENTION: Vaccine Against
TITLE OF INVENTION: Feline Rhinotracheitis
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,321
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,183
FILING DATE: July 26, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-166
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 6010703e
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 370
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE:
DESCRIPTION: Polypeptide
HYPOTHETICAL: No
ANTI-SENSE: No
ORIGINAL SOURCE:
ORGANISM: Feline herpesvirus-1
STRAIN: 1
INDIVIDUAL ISOLATE: C-27
CELL TYPE: N/A
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD: Deduced Sequence
OTHER INFORMATION: 91

US-08-911-321-10
Query Match 7.8%; Score 145; DB 3; Length 370;
Best Local Similarity 22.0%; Pred. No. 1.1e-06;
Matches 72; Conservative 52; Mismatches 136; Indels 68; Gaps 14;
QY 5 LGTALLAATLAPFGAMGIVTGNHVSARIDDDHIVVAPPE-ATLQQLQFPMFGQ-RP 62
DB 1 MSSIAFVILMAIGTVGVYRGDHSVLSLHVDTSSTSGFVIYPTLENFTIYGLIFLDDQPLP 60

QY 63 HKPYSGTVVAFRSDITNQCYQELSEERFENCSTRSSSVFVGC--KVTEYTFSSASRLTG 120
DB 61 VNNYNGTLEI--IHYNHSSCKYIVQVIEYSSCPVRNNAFRNSCHUKTSMEOYDQLSINTS 119
QY 121 PPHPFKLTIRNPRNDGMPYVIVRLDDTKBPIDVFAIQLSVQFANTAAATRGYSKASC 180
DB 120 VETGMLLTITSPKMGDGIYALRVFNHNNK-ADVFGLSVFVYSF---DFRGHRRHADE 174
QY 181 RTFG--LPTVO--LEAYLRTEESWRNWOAY---VATEATTTTSAEATTPPTVATASASELEA 234
DB 175 NLNGEILLTPSPMETIYVKV---NTPYDHMTVTQTTSNKSMESEPSNTSISC-----223
QY 235 EHFTFPWLENGVDHYEPTPANENSNTVRL-----GTMSPTLIGVTVA 278
DB 224 -----HTFQNDPNEGETLYTHLLNIAGNITYDDMDVMDGTTLKPRLL-----264
QY 279 VVSATIGLIVIVISIVTRNNCTPHEKLDIT 306
DB 265 ----DMGLNLSVTSSFKM--GNHAKMDT 286

RESULT 9
5352575-7
; Patent No. 5352575
; APPLICANT: PETROVSKIS, ERIK A.; POST, LEONARD E.; TIMMINS, JAMES G.
; TITLE OF INVENTION: PSEUDORABIES VIRUS PROTEIN
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/513,282
; FILING DATE: 20-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 100,817
; FILING DATE: 29-JUN-1987
; APPLICATION NUMBER: 886,260
; FILING DATE: 16-JUL-1986
; APPLICATION NUMBER: 784,787
; FILING DATE: 04-OCT-1985
; APPLICATION NUMBER: 801,799
; FILING DATE: 26-NOV-1985
; APPLICATION NUMBER: 844,113
; FILING DATE: 26-MAR-1986
; SEQ ID NO: 7;
; LENGTH: 350
5352575-7

Query Match 7.7%; Score 143.5; DB 6; Length 350;
Best Local Similarity 22.8%; Pred. No. 1.4e-06;
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;
QY 8 LALLAATLAPFGAM--GIVITGNHVSARIDDDHIVVAPPEATLQQLQFPMFGORPHKP 65
DB 15 LLLAALTTLAALTPTRVGGVLFRTGAAVSVHVAGSALVPGDAPNLTIDGTLLFLEGSPSN- 73
QY 66 YSGTVVAFRSDITNQCYQELSEERFENCSTRSSSVFVGC--KVTEYTFSSAS--NRL 118
DB 74 YSGRVEL--LRDPKCACTREYAAEYDLCPRVHEAFRGCLRKREPLARRASAAVEARRL 132
QY 119 TGPFPHPFKLTIRNPRNDGMPYVIVRLDDTKBPIDVFAIQLSVQFANTAAATRGIL----174
DB 133 -----LFVSRPAPPDAGSVLRVRVNGT---TDLFVLTAIV-----PPRGRPHHP 174
QY 175 --YSKASCTFGLPTVQLEAYLRTEESWRNWOAYV--ATEATTTT-----AEATPTPV 224
DB 175 TPSSADECR-----PVV-----GSWHDLSRVVDPAEDAFTTTPPIEPPEPTTPAPP 221
QY 225 TATSAS-----ELEAEHFTFPWLENGVDHYEPTPANENSNTVRLGTMSTLIGVTVA 277
DB 222 RTGATPEPRSEDEDE-----EGATTAMTPVPGILDANGTM---VLNASVSVRL 271
QY 278 AVVSATIG-----LVIVISIVTRNMCTPHRKLDIVSQDDEERSQTSRESRK 323
DB 272 AANATAGAGPGKTAWLGTPIVILLIFLGGVCAARRCARGIASTGRDPCGAARRSTR 331

RESULT 10
US-08-924-345-2
; Sequence 2, Application US/08924345
; Patent No. 6224878
; GENERAL INFORMATION:
; APPLICANT: LEUNG-PACK Patricia
; APPLICANT: LEGASTELOIS Isabelle, Christine, Marie-Andree
; APPLICANT: AUDONNET Jean-Christophe, Francis
; APPLICANT: RIVIERE Michel, Emile, Albert
; TITLE OF INVENTION: Mutants and vaccines of the Infectious
; TITLE OF INVENTION: Bovine Rhinotracheitis virus
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LARSON AND TAYLOR
; STREET: 727 SOUTH TWENTY-THIRD STREET
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,345
; FILING DATE: 04-SEP-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/199,172
; FILING DATE: 09-AUG-1994
; APPLICATION NUMBER: FR 92 07930
; FILING DATE: 26-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: SARRO, THOMAS P.
; REGISTRATION NUMBER: 19396
; REFERENCE/DOCKET NUMBER: XI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 920-7200
; TELEFAX: (703) 892-8428
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-924-345-2

Query Match 7.5%; Score 140.5; DB 3; Length 380;
Best Local Similarity 25.3%; Pred. No. 3.4e-06;
Matches 68; Conservative 33; Mismatches 125; Indels 43; Gaps 9;
QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPPEATLQL--QLFFMP 58
DB 1 MRCULLMWVLAARAAP--ARSLVYRGAVGLRAGDGVFAVHP-TDATTALGRLLIFLE 57
QY 59 GQRP-HKPYSGTVRVAFRSDITNOCYQELSEERFENCCHTRSSSVFVGCKVTEYTPFSASNR 117
DB 58 HQLPAGRYNGTVEL-LRYHAAGDCFVMLQTTAFASCPRVANNAFRSLHADTPARSER 116
QY 118 LTGP--PHFPKLTIRNPNDGSMFYIVRL-----DOTKEPIDVFAIQLSVYQFA----- 166
DB 117 RASAAVENHVLFSIARPPIDSGLYFLRVGIYGGTAGSERRRDFVPLAAFFVHSFGEGDP 176
QY 167 -----NRAATGLYSKASCTFFGL-----PTVQLEAYLRTESWR 201
DB 177 EAAATPAPRSQRPAASGLTSSASLYDRALARSFQAPPPRAPPRAARAGRRPRVDE 236
QY 202 NWQAVVATEATTTSAEATTPPVVATAS 230
DB 237 TTEVEAATRAGSAFALTTPPAGPTASPA 265

RESULT 11
US-09-175-928-2
; Sequence 2, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-175-928-2
Query Match 6.0%; Score 112; DB 4; Length 387;
Best Local Similarity 19.2%; Pred. No. 0.0036;
Matches 77; Conservative 58; Mismatches 129; Indels 138; Gaps 20;
QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPPEATLQLQLFFMPGQ 60
DB 12 LSLAGQSVVQVTI-PDGFVNVTVGSNVTLICI---YTTTVASREQLSIQWSFF----- 62
QY 61 RPKPYSGTVRVAFRSDITNOCYQELSEERFENCCHTRSSSVFVGCKVTEYTPFSASNRITG 120
DB 63 --HKMEPISYF-----SQGQVAIG-----QFKDRITG 92
QY 121 PPHP--FKLTIRNPNDGSMFYIVRLDDTKETPIDVFAIQLSVVQVFANTATRLGLYSKA 178
DB 93 SNDPGNASITISHMQPADSGIY-----ICDNNPPDFLGQNGILNVSLVKP-----SKP 143
QY 179 SCRTFGLP-----TVQLRAYLRTEESWRNQAVVATEATTTSAEATTPPVVATASSELEA 234
DB 144 LCSVQGRPETHGTISLSCL-----SALGTPSPV----- 171
QY 235 EHFTFPWLENGVDHYEPTPANENSNVTV-----RLG----- 265
DB 172 ----YYW--HKLEGRDIVPKENFNPPTTGILVIGLNTFEGQYVQCTAINRLGNSSCED 225
QY 266 -TMSPTLIGTVVAUVSATIGLVIVISIVTRNMCTPHKLDTVSODDEERSOTRESRKF 324
DB 226 LTSSHPEVGIIVGALIGSLVGAIIISVV-----CFARNKAKAKAK--ERNSTIAE---L 276
QY 325 GPMVACEINKGADQDS---ELVELVAIVNPSAL---SSPDSIK 361
DB 277 EPMT--KINPRGESEAMPREDATQLEVTLPSSIHETGPDITQ 316
RESULT 12
US-08-190-802A-63
; Sequence 63, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: TUPl, Fig. 46
US-08-190-802A-63

Query Match 5.8%; Score 108; DB 1; Length 713;
Best Local Similarity 27.3%; Pred. No. 0.025; Indels 46; Gaps 11;
Matches 62; Conservative 30; Mismatches 89; Indels 46; Gaps 11;

QY 60 QRPKPYSGTVRVAFRSDITNQCQELSEERFENCSTRSSVFGCKVTEYTFSS---AS 115
DB 86 QRDQIASLTVCQQQQQQQQVQHLQQQQ-QQLAASASVPVAQPPATTSATAPAA 144
QY 116 NRLTGPPHPFKLTIRNPRNDSGMFYVIVRLDDTKPEI-----D 154
DB 145 NTTGSPSAF--PVQASRPNLVGS-----QLPTTLTPVSSNAQQQLPQQQLQQQ 197
QY 155 VFAIQLSVYQFANTAATGLYSKASCRFTGLPTVOL-EAYLRTESWRNQAIV-----A 208
DB 198 QPPQVSVAPLSNTAING---SPTSKEITTLPSVKAPESTLKETEPENNNTSKINDTGA 254
QY 209 TEATTTTSABAT--TPTPTVATSAEAEHFTFPWLENGVDHYEPTP 253
DB 255 TTATTTTATETIKPEEDATPAS-LHQDHYLVFPYNQR-ANHSKPIP 299

RESULT 13
US-08-477-346-63
Sequence 63, Application US/08477346
Patent No. 6262023
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/477,346
APPLICATION NUMBER: US/08/477,346
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,072
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: TUPl, Fig. 46
US-08-477-346-63

Query Match 5.8%; Score 108; DB 3; Length 713;
Best Local Similarity 27.3%; Pred. No. 0.025; Indels 46; Gaps 11;
Matches 62; Conservative 30; Mismatches 89; Indels 46; Gaps 11;

QY 60 QRPKPYSGTVRVAFRSDITNQCQELSEERFENCSTRSSVFGCKVTEYTFSS---AS 115
DB 86 QRDQIASLTVCQQQQQQQQVQHLQQQQ-QQLAASASVPVAQPPATTSATAPAA 144
QY 116 NRLTGPPHPFKLTIRNPRNDSGMFYVIVRLDDTKPEI-----D 154
DB 145 NTTGSPSAF--PVQASRPNLVGS-----QLPTTLTPVSSNAQQQLPQQQLQQQ 197
QY 155 VFAIQLSVYQFANTAATGLYSKASCRFTGLPTVOL-EAYLRTESWRNQAIV-----A 208
DB 198 QPPQVSVAPLSNTAING---SPTSKEITTLPSVKAPESTLKETEPENNNTSKINDTGA 254
QY 209 TEATTTTSABAT--TPTPTVATSAEAEHFTFPWLENGVDHYEPTP 253
DB 255 TTATTTTATETIKPEEDATPAS-LHQDHYLVFPYNQR-ANHSKPIP 299

RESULT 14
US-08-473-089-63
Sequence 63, Application US/08473089
Patent No. 6342368
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,089

FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.22
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: TUP1, Fig. 46
US-08-473-089-63

Query Match 5.8%; Score 108; DB 4; Length 713;
Best Local Similarity 27.3%; Pred. No. 0.025;
Matches 62; Conservative 30; Mismatches 89; Indels 46; Gaps 11;

Qy 60 QRPKPYSGTVRAFRSDITNOCYQELSEERFENCNTHRSSSVFVGCKVTEYTPS-----AS 115
Db 86 QRDHQIASLTVOQQQQQQQQQQVQVHLQQQQ-QQLAASASVPAQPPATTSATATPAA 144

Qy 116 NRLTGPPEPKLTIRNRPNDSGMFYVIVRLDDTKPEI-----D 154
Db 145 NNTTGSAP--PVQASRPNLVGS-----QLPTTLPVVSSNAQQQLPQQQLQQQLQQQ 197

Qy 155 VFAIQLSVYQFANTAATRGLYSKASCTFGLPTVOL-EAYLRTEESWRNQAIV-----A 208
Db 198 QPPQVSVAPLSNTAING---SPTSKEITTLPSVKAPESTLKETEPENNTSKINDTGSA 254

Qy 209 TEATTTSAEAT--TPTPVATSAEAEHFTFPWLENGVDHYEPTP 253
Db 255 TTATTTTATEIKPEEDATPAS-LHQDHYLVFPYNQ-ANHSKPIP 299

Search completed: March 8, 2004, 02:46:02
Job time : 32 secs

RESULT 15
US-08-487-072A-63
Sequence 63, Application US/08487072A
Patent No. 6423684
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrisson & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,072A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:

Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQ 60
DB 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQ 60

QY 61 RHPKPYSGTVRAFRSDITNOCYQELSEERFENCSTRSSSVFVGCKVTEYTFASNRLTG 120
DB 61 RHPKPYSGTVRAFRSDITNOCYQELSEERFENCSTRSSSVFVGCKVTEYTFASNRLTG 120

QY 121 PPHPKLITIRNPRNDGSMFYVIVRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
DB 121 PPHPKLITIRNPRNDGSMFYVIVRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180

QY 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPVTATSAELEAEHFTFP 240
DB 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPVTATSAELEAEHFTFP 240

QY 241 WLENGVDHYEPTPANENSNTVRLGTMSPPTLIGTVAAVVSATIGLVIVISIVTRNMCTP 300
DB 241 WLENGVDHYEPTPANENSNTVRLGTMSPPTLIGTVAAVVSATIGLVIVISIVTRNMCTP 300

QY 301 HRKLDTVSQDDERSQTRRESKFGPMVACINKGADQDSELVELVAIVNPSALSSPDSI 360
DB 301 HRKLDTVSQDDERSQTRRESKFGPMVACINKGADQDSELVELVAIVNPSALSSPDSI 360

QY 361 KM 362
DB 361 KM 362

RESULT 2

US-09-994-064-7
; Sequence 7, Application US/09994064
; Publication No. US20030082788A1
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/994,064
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/468,190
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-994-064-7

Query Match 100.0%; Score 1863; DB 10; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.7e-179;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQ 60
DB 1 MASLLGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQ 60

QY 61 RHPKPYSGTVRAFRSDITNOCYQELSEERFENCSTRSSSVFVGCKVTEYTFASNRLTG 120
DB 61 RHPKPYSGTVRAFRSDITNOCYQELSEERFENCSTRSSSVFVGCKVTEYTFASNRLTG 120

QY 121 PPHPKLITIRNPRNDGSMFYVIVRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
DB 121 PPHPKLITIRNPRNDGSMFYVIVRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180

QY 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPVTATSAELEAEHFTFP 240
DB 181 RTFGLPTVQLEAYLRTEESWRNQAYVATEATTTSAEATTPVTATSAELEAEHFTFP 240

QY 241 WLENGVDHYEPTPANENSNTVRLGTMSPPTLIGTVAAVVSATIGLVIVISIVTRNMCTP 300
DB 241 WLENGVDHYEPTPANENSNTVRLGTMSPPTLIGTVAAVVSATIGLVIVISIVTRNMCTP 300

QY 301 HRKLDTVSQDDERSQTRRESKFGPMVACINKGADQDSELVELVAIVNPSALSSPDSI 360
DB 301 HRKLDTVSQDDERSQTRRESKFGPMVACINKGADQDSELVELVAIVNPSALSSPDSI 360

QY 361 KM 362
DB 361 KM 362

RESULT 3

US-09-994-064-68
; Sequence 68, Application US/09994064
; Publication No. US20030082788A1
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/994,064
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/468,190
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:

LENGTH: 362 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-994-064-68

Query Match 100.0%; Score 1863; DB 10; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.7e-179;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLIGTALLAATAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQLOLFMPGQ 60
DB 1 MASLIGTALLAATAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQLOLFMPGQ 60
QY 61 RHKPYSGTVRVAFRSDITNOCYQELSEERFENCNTHRSSSVFVCKTVEYTSASNELTG 120
DB 61 RHKPYSGTVRVAFRSDITNOCYQELSEERFENCNTHRSSSVFVCKTVEYTSASNELTG 120
QY 121 PHPPFKLIRNPRNDSGMFYIVRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
DB 121 PHPPFKLIRNPRNDSGMFYIVRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASC 180
QY 181 RTFGLPTVQLEAYLRTESWRNQAYVATEATTSABATTPPVTAASASELEAEHFTFP 240
DB 181 RTFGLPTVQLEAYLRTESWRNQAYVATEATTSABATTPPVTAASASELEAEHFTFP 240
QY 241 WLENGVDHYEPTPANENSVTVRLGTMSPTLIGTVAAVGSATIGLVIVISIVTRNCTP 300
DB 241 WLENGVDHYEPTPANENSVTVRLGTMSPTLIGTVAAVGSATIGLVIVISIVTRNCTP 300
QY 301 HKLQTVSDDEERSQTRERKFGPMVACINKGADQDSBELVELVAIVNPSALSPPDSI 360
DB 301 HKLQTVSDDEERSQTRERKFGPMVACINKGADQDSBELVELVAIVNPSALSPPDSI 360
QY 361 KM 362
DB 361 KM 362

RESULT 4
US-10-156-275-56
Sequence 56, Application US/10156275
Publication No. US20030049844A1
GENERAL INFORMATION:
APPLICANT: Haanes, Elizabeth J.
TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/156,275
FILING DATE: 28-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/399,118
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/680,726
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-46-C1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 364 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-10-156-275-56

Query Match 7.9%; Score 148; DB 14; Length 364;
Best Local Similarity 27.3%; Pred. No. 4.9e-06;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LIGTLALLAATAPFGAMGIVITGNHVSARIDDDHIVIVAPRPEATIQLOLFMPGQ-P 62
DB 16 LLIIMFLFLEFLYGVNGFVKQYISMFINTSSGFSIFPDDKFIYVSGRLFLDDQHL 75
QY 63 HKPYSGTVRVAFRSDITNOCYQELSEERFENCNTHRSSSVFVGC--KTEYTFASNELTG 120
DB 76 VNNYSGTIEFI--HFNNSCYTVYQTIYFSCPRIFNNAFRSLKVKYKHEHESQLRINS 132
QY 121 PHPPFKLIRNPRNDSGMFYIVRLDDTKPIDVFAIQLSVYQF 165
DB 133 IENGVLLEITNPKENDSGVIRVQLENNK--TDVGIAPFIYF 175

RESULT 5
US-10-125-692-10
Sequence 10, Application US/10125692
Publication No. US20030044429A1
GENERAL INFORMATION:
APPLICANT: Aderem, Alan
APPLICANT: Hayashi, Fumitaka
APPLICANT: Smith, Kelly D.
APPLICANT: Underhill, David M.
APPLICANT: Ozinsky, Adrian
TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods
FILE REFERENCE: P-IS 5155
CURRENT APPLICATION NUMBER: US/10/125,692
CURRENT FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: US 60/285,477
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 612
TYPE: PRT
ORGANISM: Homo sapiens
US-10-125-692-10

Query Match 6.2%; Score 115; DB 14; Length 612;
Best Local Similarity 20.6%; Pred. No. 0.024;
Matches 76; Conservative 54; Mismatches 133; Indels 106; Gaps 16;

QY 19 GAGGIVITGNHVSARIDDDHIVIVAPRPEATIQLOLF--FMFGQPHKPYSGTVRVAFR 75
DB 93 GSFPFLII--NKLKWDSTQVCELENRKE-EVELWFKVTFSGTSLQGSLLTLDN 149
QY 76 SDITNOCYQELSEERFENCNTHRSSSVFVGCYKTEYTFASNRLTGPFPKLTIRNPREN 135
DB 150 SKVSN---PLTE-----CKKKGKGVSGSKV-----LSMSNLRVQ 181
QY 136 DSGMFYIVRLDDTKPIDVFAIQLSVYQFANTAATRGLYSKASCRTFGLPTVQLEAYLR 195
DB 182 DSDFNCTVILDQKK--NWFGLSVLGFQSTAIT-AKSEGESAEFFPLNFAE---- 233
QY 196 TEESWRN--WQAYVATEATTTSAAETTPVTATSAELEAEHFTFPWLENGVDHYEPTP 253
DB 234 -ENCGELMWKA-----EKDSFFQFQWISFSIKKKEVS 265

Db 198 LSLTPSLP--BEATGLTTPESKEGPIITAETVLPDSWSVSTSDTAVLLTSKES 255
Qy 292 IVTRMCTPH-----RKLDTVSQDDEERSQT-----RRSRKFGPMVACEINKGADQDSELV 343
Db 256 KWDLPSTSHVSMWKTSDSVSPQSGASDTAVPEQNKTKTKGQMDG--IPMSMKNMEMPIS 313
Qy 344 ELVAIVNPS 352
Db 314 QLLMIAPS 322

RESULT 11

US-10-188-012-33
; Sequence 33, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 378
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(378)
; OTHER INFORMATION: TIM-4, allele 1
US-10-188-012-33

Query Match 5.5%; Score 102.5; DB 14; Length 378;
Best Local Similarity 23.6%; Pred. No. 0.21;
Matches 73; Conservative 36; Mismatches 133; Indels 67; Gaps 13;
Qy 65 PYSGTVRVAFRSDITNQCQYQELSEERFENCTHRSSSVFVGCKVTEYTFPSASRLTG--PP 122
Db 60 PYSCKEALIRTD-----GMRVTSRK-SAKYRLOQTIPR 92
Qy 123 HPFKLTIRNPRNDGSMFYVIVRL-----DDTKEPIDVFAIOLSVYQFANTAATRGLYSKA 178
Db 93 GDSVLTILNPESDSGVYCCRIEVPGFNDVK-----INVRNL-QRASTTTHRTATTTT 146
Qy 179 SCRTFGLPTVQLEAYLRTEESWRNMQAVATEATTSAEATTPP-----VTATSASE 231
Db 147 RRTTTSPT-----TTRQMTTTPAALPTTVTT-PDLTTGTPLQMTTIAVFTTANTC 197
Qy 232 LEAEHFTFPWLENGVDHYEPTPANENSVTVRLGTMGPTLIGTVAAVWATIGLVIVIS 291
Db 198 LSLTPSLP--BEATGLTTPESKEGPIITAETVLPDSWSVSTSDTAVLLTSKES 255
Qy 292 IVTRMCTPH-----RKLDTVSQDDEERSQT-----RRSRKFGPMVACEINKGADQDSELV 343
Db 256 KWDLPSTSHVSMWKTSDSVSPQSGASDTAVPEQNKTKTKGQMDG--IPMSMKNMEMPIS 313
Qy 344 ELVAIVNPS 352
Db 314 QLLMIAPS 322

RESULT 12

US-09-813-153-138
; Sequence 138, Application US/09813153

; Publication No. US20030045459A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 67 Human secreted proteins
; FILE REFERENCE: PZ023
; CURRENT APPLICATION NUMBER: US/09/813,153
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/363,044
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,167
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,162
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,161
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,170
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 138
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (379)
; OTHER INFORMATION: Xaa equals stop translation
US-09-813-153-138

Query Match 5.5%; Score 102.5; DB 10; Length 379;
Best Local Similarity 23.6%; Pred. No. 0.21;
Matches 73; Conservative 36; Mismatches 133; Indels 67; Gaps 13;
Qy 65 PYSGTVRVAFRSDITNQCQYQELSEERFENCTHRSSSVFVGCKVTEYTFPSASRLTG--PP 122
Db 60 PYSCKEALIRTD-----GMRVTSRK-SAKYRLOQTIPR 92
Qy 123 HPFKLTIRNPRNDGSMFYVIVRL-----DDTKEPIDVFAIOLSVYQFANTAATRGLYSKA 178
Db 93 GDSVLTILNPESDSGVYCCRIEVPGFNDVK-----INVRNL-QRASTTTHRTATTTT 146
Qy 179 SCRTFGLPTVQLEAYLRTEESWRNMQAVATEATTSAEATTPP-----VTATSASE 231
Db 147 RRTTTSPT-----TTRQMTTTPAALPTTVTT-PDLTTGTPLQMTTIAVFTTANTC 197
Qy 232 LEAEHFTFPWLENGVDHYEPTPANENSVTVRLGTMGPTLIGTVAAVWATIGLVIVIS 291
Db 198 LSLTPSLP--BEATGLTTPESKEGPIITAETVLPDSWSVSTSDTAVLLTSKES 255
Qy 292 IVTRMCTPH-----RKLDTVSQDDEERSQT-----RRSRKFGPMVACEINKGADQDSELV 343
Db 256 KWDLPSTSHVSMWKTSDSVSPQSGASDTAVPEQNKTKTKGQMDG--IPMSMKNMEMPIS 313
Qy 344 ELVAIVNPS 352
Db 314 QLLMIAPS 322

RESULT 13

US-09-864-408A-5760
; Sequence 5760, Application US/09864408A
; Publication No. US2004000947A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.

```

; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5760
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-5760

Query Match      5.5%; Score 102; DB 11; Length 234;
Best Local Similarity 26.9%; Pred. No. 0.11;
Matches 61; Conservative 30; Mismatches 90; Indels 46; Gaps 11;

QY 60 QRPKPYSGTVRVAFRSIDINQCYQELSEERFENCNTHRSSVFGCKVTEYTFSS-----AS 115
Db 15 QRDQIASLTVQQRQQRQQRQVQHLLQQQ-QQLAASASVPVQAQPPATTSATATPAA 73
QY 116 NRLTGPFPKLTIRNRPNDNGMFYVIVRLDDTKPEI-----D 154
Db 74 NITGSPSAP--PVQASRPNLVGS-----QLPTTLEWSSNAQQQLPQQOQQOQQOQQ 126
QY 155 VPAQLSVQFQANTAAATGELYSKASCTFGLPTVQL-EAYLREESWRWQAVY-----A 208
Db 127 QPPQVQVAPLSNTAING---SPTSKEITLPSVKAPESTLKTETEPENNNTSKINDTGA 183
QY 209 TEATTSAAEAT--TPTVPTATSAEAEAEHFTFPLENGVDHVEPTP 253
Db 184 TTAITTTATEIKPKEDATPAS-LAQDHYLVFYNQR-AHNSKPIP 228

RESULT 14
US-10-295-027-1308
; Sequence 1308, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08

```

```

; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1308
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1308

Query Match      5.4%; Score 101; DB 15; Length 413;
Best Local Similarity 19.7%; Pred. No. 0.34;
Matches 80; Conservative 61; Mismatches 144; Indels 122; Gaps 21;

QY 1 MASLGLTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPPEATIQQLFFMPGQ 60
Db 12 LSLAGQSVVQVTI-PDGFVNVTGVSNTLICI---YTTTVAEREQLSIQWSPFH---K 64
QY 61 RPHKPYSGTVRVAFRSIDINQCYQELSEERFENCNTHRSSVFGCKVTEYTFSSAS 115
Db 65 KEMETISSPWEEGKWPDEAVKGLDGOAELQIVF---SQGQAVATG-----QPK 113
QY 116 NRLTGPFPKLTIRNRPNDNGMFYVIVRLDDTKPEIDVFAQLSVQFQANTAAATG 173
Db 114 DRITGNDPAGNATITISHMQPADSGIY-----ICDVNNPDPFLGQNGILNVSLVKP-- 166
QY 174 LYSKASCTFGLP-----TVQLEAYLREESWRWQAVYATEATTSAAETPTPTATSA 229
Db 167 --SKPLCSVQGRFPGTHTISLCL-----SALGTSPV----- 197
QY 230 SELEAEHFTFPLENGVDHYPPTANENSVTV-----RLG-- 265
Db 198 -----YV--HKLEGRDIPVKENFNPTTGLIVGLNLTNFEQGYVQCTAINRLGNS 246
QY 266 -----TMSPTLIGTVAAVVSATIGLVIVISIVTRNNCTPHRKLDTVQSDDEERSQTRR 319
Db 247 SCEIDLTSSHPEVGIIVGALIGSLVGAALIIISV---CFARNKAKAKAK--ERNSKTIA 300
QY 320 ESRKFGPMVACINKGAQODS---ELVELVAIVNPSAL--SSPDSIK 361
Db 301 E---LEPMT--KINPRGSEAMPREDATQLEVTLPSSIHETGPDTIQ 342

RESULT 15
US-10-282-122A-60126
; Sequence 60126, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

```

```
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60126
; LENGTH: 961
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-60126

Query Match      5.3%; Score 99; DB 12; Length 961;
Best Local Similarity 18.5%; Pred. No. 1.9;
Matches 48; Conservative 45; Mismatches 95; Indels 72; Gaps 9;

QY      20  AMGIVITGNHVSARIDDDHIVIVAPRPEATLQLLFMPGQRPKPYSGTVRVAFRSDIT 79
Db      427  ALDVVNIADQYDPQAIKDRLAMTEQ-----NARIWISPOEPHNKTAYFVDAFYQVD-- 479

QY      80  NCYOELSEBERFENCETHRSSSVFVCKV-----TEYTFASNRLTGPPHP----- 124
Db      480  -----KISEQTFADWQHKSOATLQLPALNPYIPDDFTLIKSDKAW--PHFQLILDBFTL 532

QY      125  -----FKLTIRNPRNDSCGMFVIVIVELDDTKEPIDVFAIQLSVYQFAN 167
Db      533  RVVYAPSOYFASEPKADISLVRNPQAMDSARRQWFMALND-----YIAGIALDDQLSN 585

QY      168  -----TAATRGLYSKASCRTFGLPTVQLE-----AYLRTEESW---RNWQAYVA 208
Db      586  QAAVGISFSTGANGNGLMVNANGYQHLPALFSDLLQGYFSYPTTEEQLEQAKSWYACMM 645

QY      209  TEATTTSAEATTTFTPVVATS 228
Db      646  DSAEKGYDQAIMPIQWYS 665
```

Search completed: March 17, 2004, 23:21:55
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 02:31:35 ; Search time 42 Seconds
(without alignments)
829.079 Million cell updates/sec

Title: US-09-993-777-7

Perfect score: 1863
Sequence: 1 MASLLGLTALLATLAPFGA.....VELVAIYNPSALSPDSIKM 362

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*

- 1: PIR1.*
- 2: PIR2.*
- 3: PIR3.*
- 4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	8.8	420	2 T42616	envelope protein -
2	162.5	8.7	354	1 VGBE67	glycoprotein D pre
3	155	8.3	353	1 C46113	glycoprotein D pre
4	153.5	8.2	424	1 VGBE89	glycoprotein gp63
5	143.5	7.7	350	1 VGBE83	glycoprotein gp63
6	140.5	7.5	380	2 S35785	glycoprotein I - b
7	125	6.7	356	2 JQ2352	glycoprotein I - t
8	124.5	6.7	683	2 T03146	probable glycoprot
9	115	6.2	457	1 RWM574	T-cell surface gly
10	113.5	6.1	390	1 Q0BE77	glycoprotein I pre
11	111.5	6.0	149	2 A61162	glycoprotein I - e
12	108.5	5.8	355	2 C39725	hypothetical prote
13	108	5.8	713	2 JN0133	WD-40 repeat regul
14	107	5.7	372	1 Q0BE88	glycoprotein I pre
15	99	5.3	372	2 F43674	US7 protein - huma
16	99	5.3	1036	2 S73601	protein p200 - Myc
17	98.5	5.3	1140	2 S73786	hypothetical prote
18	98	5.3	503	2 B38745	cell adhesion mole
19	98	5.3	507	2 S64507	probable membrane
20	97	5.2	1220	2 S48387	SIN1 protein - yea
21	96.5	5.2	814	1 A39752	fibroblast growth
22	95.5	5.1	797	1 VGBEX1	glycoprotein X pre
23	95.5	5.1	867	2 T45463	membrane glycoprot
24	95	5.1	1777	2 T34369	hypothetical prote
25	94	5.0	796	2 T21460	hypothetical prote
26	93.5	5.0	645	2 T29818	hypothetical prote
27	93.5	5.0	907	1 Q0BE21	membrane antigen g
28	93.5	5.0	3020	2 A43932	mucin 2 precursor,
29	93	5.0	1051	2 S55259	TIF1 protein - mou

30	92.5	5.0	1778	2 AF1116	internalin protein
31	92	4.9	1044	2 F95375	probable drug resi
32	91.5	4.9	510	2 T47374	hypothetical prote
33	91	4.9	365	2 A34424	CD44 membrane gly
34	91	4.9	836	2 S49940	cell division cont
35	90.5	4.9	547	1 S28904	intercellular adhe
36	90.5	4.9	785	2 T37329	probable oligopept
37	90.5	4.9	786	2 T19017	hypothetical prote
38	90.5	4.9	1429	2 T41699	C2-domain family p
39	90	4.8	233	2 C84193	hypothetical prote
40	90	4.8	798	2 T34248	hypothetical prote
41	89.5	4.8	826	2 G90283	hypothetical prote
42	89.5	4.8	866	2 T45462	membrane glycoprot
43	89.5	4.8	992	2 T38817	hypothetical prote
44	89	4.8	629	2 AE1525	probable peptidogl
45	89	4.8	678	2 S54308	DNA binding protei

ALIGNMENTS

RESULT 1

T42616

envelope protein - equine herpesvirus 4 (strain NS80S67)

C:Species: equine herpesvirus 4

A:Variety: strain NS80S67

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C:Accession: T42616

R:Telford, B.A.; Watson, M.S.; Perry, J.; Cullinan, A.A.; Davison, A.J.

J. Gen. Virol. 79, 1197-1203, 1998

A:Title: The DNA sequence of equine herpesvirus-4.

A:Reference number: Z22173; MUID:98264497; PMID:9603335

A:Accession: T42616

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-420 <TSL>

A:Cross-references: EMBL:AF030027; NID:G2605950; PIDN:AAC59593.1; PID:G2606021

A:Experimental source: strain NS80S67

C:Genetics:

A:Note: 73

C:Superfamily: pseudorabies virus glycoprotein gp63

Query Match 8.8%; Score 164; DB 2; Length 420;
Best Local Similarity 25.5%; Pred. No. 9.6e-06;
Matches 96; Conservative 50; Mismatches 143; Indels 88; Gaps 21;

Qy	10	LLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQIQLFEMFGQR-PHKPYS	67
Db	13	LLAISMCL---ATAIIVRGERHSMYLNASEFAVYPKDKSLVVVGHMLFDGQRLPTNTS	69
Qy	68	GTVRVAFRSDITNQCQELSEERFENCNTHRSSSVFVGC--KVTEYTFSSANRLTGPPHP	125
Db	70	GLIEL-IHNVSRGCYSVIQITISYSCPRVANNAFRSLHKTSHNHQDYFHVNTSVETNV	128
Qy	126	KLITRNPRNDGMFVIVRLDDTKPIDVFAIQLSVQF-ANTATRGLYSKASCRTPG	184
Db	129	LLNITRQPADSGAYILRLVNLHA-PTADVGVSAFVYDLSQNTPEPVPVTAKEPSNVET	187
Qy	185	LPVQLAEYLRTESRNQAVYVATEATTTSABAT---TPTPVATASASELEAHEHTPW	241
Db	188	RTAPAPANTSTK-----TGSNTSSOSTWLYTPTP-----RPA	221
Qy	242	LENGVDHYETPANEN--SNVTVRL-----GTMSPT-----LIGVTVAASVSGATIG	285
Db	222	LET---HLTAPANETVWSGDTAMLCHGRFPSTAVTIYMHLLGLT-----GNLPE	269
Qy	286	--LVIVISIVTRNMCTPHRKLDTVS-----QDDEERSQTRRESR-KFGPMV-----ACEIN	333
Db	270	DVLLIEDSEILR---TPPKPQTTSSRTGDDFKQTNSTSPKSRNKIVAMVVIPTACVLM	326
Qy	334	KGADQDSSELVELVAIYN	350
Db	327	-----LLLVVVGAIIN	337

```
F;21-33/Product: glycoprotein D #status predicted <GPD>
F;275-293/Domain: transmembrane #status predicted <TMN>
F;40-75,84,122-138,227,252/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

Query Match 8.3%; Score 155; DB 1; Length 353;
Best Local Similarity 20.2%; Pred. No. 4.2e-05;
Matches 65; Conservative 66; Mismatches 151; Indels

18	FG--	AMGIVITGNHVSARIDDD--	HIVIVAPREATIQLOFPWGPQR--	HKPYSGVTR	71
QY	:	:	:	:	:
19	:	:	:	:	:
20	FGIQCAAAI	IRGNHVISLYNSSATSFLKGNNDASIRGRFLIGDQFPVTNTYNTVTE			79
DB	:	:	:	:	:
72	VAFRSDI	INQCVELSEERFENCETHRSSVFGCKV--	TEYTFASNRLLTGPHFPFKLTIR		131
QY	:	:	:	:	:
80	L--LHVNTQT	CLQPLRYVMYGECPRIRTGAIACRKGWSHEYATQTD	DNVFIIFKQD		131
DB	:	:	:	:	:
131	NPRNDSGMFFYVRLDDTKGPI	VDVAFQLSVYQFANTAATRGYSKASCTFGLPVTQOL			191
QY	:	:	:	:	:
139	NTKVEDAGI	YLLVQLDYT--SLGDIFFVSLNVYKQDTSNEDVNY-----	FFPYVS		181
DB	:	:	:	:	:

[illegible]

RESULT 4

VG8EE9
 glycoprotein gp63 precursor - equine herpesvirus 1
 N;Alternate names: glycoprotein I precursor; hypothetical 46K protein; ORF2 protein
 C;Species: equine herpesvirus 1
 A;Note: host Equus caballus (domestic horse)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C;Accession: C36846; JQ0998; A36803
 R;Audonnet, J.C.; Winslow, J.; Allien, G.; Paoletti, E.
 J. Gen. Virol. 71, 2969-2978, 1990
 A;Title: Equine herpesvirus type 1 unique short fragment
 A;Reference number: A36846; MUID:91108393; PMID:2177089

A/Accession: G36746
A/Molecule type: DNA
A/Residues: 1-424 <AUG>
A/Cross-references: NID:g330791; PIDN:RAB02508.1; PID:g330864
A/Residuals: 1-424 <AUG>
A/Experimental source: strain Kentucky D
A/R/Eiton, D.M.; Halliburton, I.W.; Killington, R.A.; Meredith, D.M.; Bonat
Gene 101, 203-208, 1991
A/Title: Sequence analysis of the 4.7-kb BamHI-EcoRI fragment of the equi
A/Reference number: JQ0998; MUID:91276272; PMID:1647359
A/Accession: JQ0998
A/Molecule type: DNA
A/Residues: 1-424 <ELT>
A/Cross-references: GB:M36299; NID:g330787; PIDN:AAA66547.1; PID:g330788
A/R/Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A/Description: The DNA sequence of equine herpesvirus-1.
A/Reference number: A36805

A;Molecule type: DNA
A;Residues: 1-424 <TEL>
A;Cross-references: GB:M36299; NID:g330787; PID:g330788
B;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A;Description: the DNA sequence of equine herpesvirus-1.
A;Reference number: A36805
A;Accession: A36803
A;Molecule type: DNA
A;Residues: 1-424 <TEL>
A;Cross-references: GB:M86664; NID:g330791; PID:AA02508.1; PID:g330864
A;Experimental source: strain Ab4p
B;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A;Title: The DNA sequence of equine herpesvirus-1.
A;Reference number: A41831; MUID:92295566; PMID:1318606
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given

RESULT 2
VGBE67

glycoprotein D precursor - human herpesvirus 3
 N:Alternate names: glycoprotein IV
 C:Species: human herpesvirus 3, varicella-zoster virus
 C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
 C:Accession: F27345
 J:R.Davidson, A.J.; Scott, J.E.
 J:Gen. Virol. 67, 1759-1816, 1986
 A:Title: The complete DNA sequence of varicella-zoster virus.
 A:Reference number: A27345; MUID:86306857; PMID:3018124
 A:Accession: F27345
 A:Molecule type: DNA
 A:Residues: 1-354 <DAV>

A: Cross-references: EMBL:X04370: NID:q59989; PIDN:CAA27950.1; PID:g60056

A:Gene: 67
C:Superfamily: herpesvirus glycoprotein D
K:Keywords: glycoprotein; phosphoprotein; transmembrane protein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-354/Product: glycoprotein D #status predicted <CPD>
F:18-354/Product: glycoprotein D #status predicted <CPD>
F:279-295/Domain: transmembrane #status predicted <TMN>
F:279-295/Domain: transmembrane #status predicted <TMN>
F:33-47-67/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match	8.7%	Score 162.5;	DB 1;	Length 354;
Best Local Similarity	22.4%;	Pred. No. 1e-05;		
Matches	72: Conservative	53: Mismatches 111;	Indels	85: Gaps 15;

23 IVITGNHVSARIDDDHIVIVAPRE--ATIQQLFFMPGQRP-HKPYSGTVRVAFRSDI 78
 21 LIFKGDHVSIVNSSLTSLIPMONDNYEIKGOLVFGEOLPTGYNYSGLTELLY-ADT 79

79 TNQCYQLSEERFENCITRSSSVVGCKVT--EYTPSASRLTGPPHPFKLTIRNPRPN 135
::
80 VAECEPSGVNTVNGCPRTPTSAITSRYKHSWHYGNSTDRISTPDAGWMLKITPGGIN 139

QY
136 DSGMFYIVRLDTEKPIDVFAIQISVY-----QPANTAATRGLYSKAS 179

| : : : | | : : | | : | : |
| : : : | | : : | | : | : |

DH
140 DACWAVI VPI DSPPS-TRGFTIGWNVYAGSHNHGVITYTSPIONGVSTRALFOOA- 197

180 CRTGFLPT-----VQLEAYLRTEES--ERNQAVVATEATTTSABATTTPVT 225

198 -PCTVDAAPKQSGSTSEFMIDLPAGKSTLENNPMTIH-EDVVTTEKSVVVEG----- 248

Qy 226 ATSA SELEAEHFTFPWLENGVDHYEPTPAN---ENS-----NVYRLGTMSPFLIGTVV 276

QY 277 AAVSATIGLVIVISVTRNM 297
: : : ||||| :
282 SIAIT TAAVIVIVISVTRST 302

RESULT 3

glycoprotein D precursor - cercopithecine herpesvirus 9 (strain DHV)
 N;Alternate names: membrane glycoprotein 1
 C;Species: cercopithecine herpesvirus 9

C;Accession: C46113
R;Fletcher III, T.M.; Gray, W.L.
Virology 193, 762-773, 1993

A;Reference number: A46113; MUID:93212509; PMID:8384754
A;Accession: C46113
A;Molecule type: DNA

A; Cross-references: GB:I07067; NID:g310715; PIDN:AAA47888.1; PID:g310718
C; Superfamily: herpesvirus glycoprotein D
C; Keywords: glycoprotein; transmembrane protein

RESULT 3

C46113
glycoprotein D precursor - cercopithesine herpesvirus 9 (strain DHV)
N;Alternate names: membrane glycoprotein 1
C;Species: cercopithesine herpesvirus 9
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C;Accession: C46113
R;Fletcher III, T.M.; Gray, W.L.
Virology 193, 762-773, 1993
A;Title: DNA sequence and genetic organization of the unique short (Us) region of the si
A;Reference number: A46113; MUID:93212509; PMID:8384754
A;Accession: C46113
A;Molecule type: DNA
A;Residues: 1-353 <FILE>
A;Cross-references: GE:I07067; NID:g310715; PIDN:AAA47888.1; PID:g310718
C;Superfamily: herpesvirus glycoprotein D
C;Keywords: glycoprotein; transmembrane protein
P;1-20/Domain: signal sequence #status predicted <SIG>

C;Genetics:
A;Gene: 73
C;Superfamily: pseudorabies virus glycoprotein gp63
C;Keywords: glycoprotein; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-424/Product: glycoprotein gp63 #status predicted <MAT>
F;320-336/Domain: transmembrane #status predicted <TM1>
F;35,67,78,121,131,236,307/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.2%; Score 153.5; DB 1; Length 424;
Best Local Similarity 23.8%; Pred. No. 7.1e-05;
Matches 70; Conservative 41; Mismatches 126; Indels 57; Gaps 12;

QY 1 MASLGLTALLAATLAPFGANGIVITGNHVSARIDDDHIVIVAPRENTIQL-QLFFMPG 59
DB 1 MAKLTGMSAALLSNAICSTALIYRHEHSMYLNASSEFAVYPTQDSLVGVHLLFLDG 60
QY 60 QR-PHKYSYSGTVRVAFRSDITNOCYQELSEERFENCNTHRSSVFGVC--KVTEYTFSA 116
DB 61 QRLPTNYSGLIEL-IHYNSSVCYTVIQISVSCPRVANNAFRSLKHTSKHYDYFR 119
QY 117 RLTPGPHFKLTIRNPNDSGMFFYIVRLDDTKEPIDVFAIQLSVYQFANTAATRLGLS 176
DB 120 VNARSVEINLLNITKPOPTDSGAILRVKLDHA-PTADVFGVSFAFYVDL----- 167
QY 177 KASCRTPGLPTVQ----LEAYLRTEESWRNMQAVATEATTS---AEATTPPTVATSA 229
DB 168 KSKTVPDPMPETQTVEPTTSVSTPYDYDDVTETESTSTQOAMTSTQTP-SATWG 226
QY 230 SELEAEHFTFPWLENGVDHVEPTPANENSVTVRLG-----TMSPTL 271
DB 227 TQLTTE-----LPTNE-----TVVIGQALLCHWFQSPTRVPTL 260

RESULT 5
VGBB63
glycoprotein gp63 - suid herpesvirus 1
C;Species: suid herpesvirus 1
A;Note: host Sus scrofa domestica (domestic pig)
C;Date: 31-Mar-1998 #sequence_revision 31-Mar-1998 #text_change 16-Jul-1999
C;Accession: A29012
R;Petrovskis, E.A.; Timmins, J.G.; Post, L.E.
J. Virol. 60, 185-193, 1986
A;Title: Use of lambda-gt11 to isolate genes for two pseudorabies virus glycoproteins w/
A;Reference number: A93021; MUID:86308235; PMID:3018284
A;Accession: A29012
A;Molecule type: DNA
A;Residues: 1-350 <P>
A;Cross-references: GB:M4336; NID:G334055; PIDN:AAC35204.1; PID:G334056
A;Experimental source: strain Rice
C;Superfamily: pseudorabies virus glycoprotein gp63
C;Keywords: glycoprotein; transmembrane protein
F;11-27/Domain: transmembrane #status predicted <TM1>
F;286-308/Domain: transmembrane #status predicted <TM2>
F;56,73,153,256,262,275/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.7%; Score 143.5; DB 1; Length 350;
Best Local Similarity 22.8%; Pred. No. 0.00037;
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;

QY 8 LALLAATLAPGAM--GVITGNHVSARIDDDHIVIVAPRENTIQLQLFFMPGQPHKP 65
DB 15 LLLAALTALATPRVGVLFPGAGVSVHVAGSALVLPDGDAPNLITDGLLFLGSPSPN- 73
QY 66 YSGTVRVAFRSDITNOCYQELSEERFENCNTHRSSVFGVC--KVTEYTFSA-----NRL 118
DB 74 YSGRVEL-LRLDPRKACVTRYAAYDLCPRVHHEAFRGCLRKEEPLARRASAAVEARL 132
QY 119 TGPHPFKLTIRNPNDSGMFFYIVRLDDTKEPIDVFAIQLSVYQFANTAATRLGL----- 174
DB 133 -----LFVSRPAPPDAGSVLVRVNGI---TDLFVLTAIV-----PPRGRPHHP 174
QY 175 --YSKASCRTPGLPTVQLEAYLRTEESWRNMQAVV---ATEATTS-----AEATTPTV 224

C;Genetics:
A;Gene: 73
C;Superfamily: pseudorabies virus glycoprotein gp63
C;Keywords: glycoprotein; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-424/Product: glycoprotein gp63 #status predicted <MAT>
F;320-336/Domain: transmembrane #status predicted <TM1>
F;35,67,78,121,131,236,307/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.5%; Score 140.5; DB 2; Length 380;
Best Local Similarity 25.3%; Pred. No. 0.00072;
Matches 68; Conservative 33; Mismatches 125; Indels 43; Gaps 9;

QY 1 MASLGLTALLAATLAPFGANGIVITGNHVSARIDDDHIVIVAPRENTIQL-QLFFMP 58
DB 1 MRCLLLMWVLAARAAP--ARSLVYRGEAVGLRADGPVAFVHP-TDATLALRGSLIFLE 57
QY 59 GQRP-HKPYSGTVRVAFRSDITNOCYQELSEERFENCNTHRSSVFGVKVTEYTFSA 117
DB 58 HOLPAGRRYNGTVEL-LRYHAAGDCFVMLQTTAFASCPRVANNAFRSLHADTRPARSER 116
QY 118 LTGP--PHPKLTIRNPNDSGMFFYIVRL-----DDTKEPIDVFAIQLSVYQFA----- 166
DB 117 RASAAVENHVLFSIARPPIDPSGLYFLRVGYGTAGSERRRDRVFPPLAAVHSGEPGDP 176
QY 167 -----NTAATRGLYSKASCRTFGL-----PTVQLEAYLRTEESWR 201
DB 177 EAAATPAPSQRSPASASGLTSSASLYDRALARSQAPPPRAPPRAARAGRRPERVDE 236
QY 202 NMQAVYATEATTTSAEATTPPTVATSA 230
DB 237 TTEVEAATFAGSAPALTTPPGPTASPA 265

RESULT 7
JQ2352
glycoprotein I - turkey herpesvirus
N;Alternate names: ORF 7 protein
C;Species: turkey herpesvirus
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 07-May-1999
C;Accession: JQ2352
R;Zelnik, V.; Dartell, R.; Audonnet, J.C.; Smith, G.D.; Riviere, M.; Pastorek, J.; Ross
J. Gen. Virol. 74, 2151-2162, 1993
A;Title: The complete sequence and gene organization of the short unique region of herp
A;Reference number: JQ2346; MUID:94014999; PMID:8409940
A;Accession: JQ2352
A;Molecule type: DNA
A;Residues: 1-356 <ZEL>
C;Keywords: glycoprotein; transmembrane protein
F;5-21/Domain: transmembrane #status predicted <TM1>
F;277-293/Domain: transmembrane #status predicted <TM2>
F;154,167,250,260/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.7% Score 125; DB 2; Length 356;
Best Local Similarity 22.2% Pred. No. 0.012;
Matches 85; Conservative 54; Mismatches 174; Indels 70; Gaps 17;

QY 6 GTLALLAATLAPF---GAMGIVITGNHVSARIDDD-HIVIVAPRPEATLQLFPMPCQ 60
DB 5 GIACWCVTLILFCIIKTCQMVYRSLVSTTDQSAQIAFFGIDTWNLYGKLVFMGQ 64
QY 61 RPKPVSCTVYAFRSDITNQCQELSEERFENCNTHRSSVFGCKVTEYTSASRLTG 120
DB 65 YLEIYSGTMEI-LKWNQANRCYSIAHATYADCPITSSVFRGCRDAVYTRPHSRI-- 121
QY 121 PPHF-----FKLIRNRPNDSGMFWYVIVELDTKEPIDVFAIQLSVQYQFANTAARGLY 175
DB 122 --HPQVNGLLTIIERMEDSGIYIRTSIDFNKS-----DYARTSIFVCG 168
QY 176 SKASCRT-----FGLPTVQ---LEAYLRTESWRNW---QAYVATEATTSAEAT 220
DB 169 SSGSCSNPRQKVSDEMCIPHNVAIPEVYLTLVHGLPYGDLTLQIRKDMITT---APT 225
QY 221 PTPVATASASELEAHEFTFWLNGVDHYEPTPANENSNVTVRLGTMSPTLIGVTVAUV 280
DB 226 YRTIRRTVNEGLLTAKTSPOIDLNATNL-PLPI---SNVT---DYNMVIWRRVALRIY 278
QY 281 SATIGLVIVISVTRNMCTPHRKLDTVSQDDEERSQTRBSRKF--PMVACEINKGADQ 338
DB 279 AYLVIATIALIIVT--VCSAHR-----GSCSRRRRIYVIGNEPTTLTSITNGNQ 326
QY 339 DSELVELVAIVNPSALSPPSIK 361
DB 327 EKETXNV-----PSDISDAELLE 344

RESULT 8
T03146
probable glycoprotein A8 - alphalaphine herpesvirus 1
C:Species: alphalaphine herpesvirus 1
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
C:Accession: T03146
R:Esser, A.; Pflanz, R.; Fleckenstein, B.
J. Virol. 71, 6517-6525, 1997
A:Title: Primary structure of the alphalaphine herpesvirus 1 genome.
A:Reference number: Z14840; MUID:97404659; PMID:9261371
A:Accession: T03146
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-683 <ENS>
A:Cross-references: EMBL:AF005370; NID:g2337967; PIDN:AAC58098.1; PID:g2338014

Query Match 6.7% Score 124.5; DB 2; Length 683;
Best Local Similarity 21.9% Pred. No. 0.032;
Matches 72; Conservative 46; Mismatches 108; Indels 103; Gaps 14;

QY 11 LAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPE--ATIQQLQFPMFGQRPKPYSG 68
DB 341 ITATISPRGEIKNVITGNF--SRNANLSLAFSSKGEYAGVLQA-FEPSTRPP---PG 394
QY 69 TVRVAFRSDITNQCQELSEERFENCNTHRSSVFGCKVTEYTSASRLTGPHPKLT 128
DB 395 TVAPGILSTAN-----FETSTNKSPTYT-----PIPAKLSPTPGLTNTL 435
QY 129 IRNRPNDSGMFWYVIVELDTKEPIDVFAIQLSVQYQFANTAARGLYSKASC----- 180
DB 436 LLTAGEHNSG-----IGSTLEPLTVSVQVLTQSPSPTRDTSTLVIKLTVDPQHKTV 488
QY 181 -----RTFGLPTVOLEAYLRTESWRNWQAYVATEATTSAEATTPPVTATGASEL 232
DB 489 SPSLVTCRTSTLIVSWTHSPRSGSPK-----PQTAKTSGSALPPLITTT----- 538
QY 233 EAEHFTFWLNGVDHYEPTPAN-----ENSNVTV-----RLGTMSPTLIGV 274
DB 539 -----PIPTNTEKQSIFASSTVSVDITFTTGDDVNTVGTWSPS--I 577

QY 275 TVAAVVSATIGLVIVISVTRNMCTPHRK 303
DB 578 TQTLPITPSTSGRQIVV-----GCCTLNAR 602

RESULT 9
RWMS74
T-cell surface glycoprotein CD4 precursor - mouse
N:Alternate names: T-cell differentiation antigen L3T4; T-cell surface antigen T4/Leu 3
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A02110; A26038; A39893; A39955; I54564; I69018; A47642
R:Tourville, B.; Gorman, S.D.; Field, E.H.; Hunkapiller, T.; Parnes, J.R.
Science 234, 610-614, 1986
A:Title: Isolation and sequence of L3T4 complementary DNA clones: expression in T cells
A:Reference number: A02110; MUID:87018845; PMID:3094146
A:Accession: A02110
A:Molecule type: mRNA
A:Residues: 1-457 <TOU>
A:Cross-references: GB:M13816; NID:g192070; PIDN:AAA37267.1; PID:g309112
R:Rittman, D.R.; Gettner, S.N.
Nature 325, 453-455, 1987
A:Title: Unusual intron in the immunoglobulin domain of the newly isolated murine CD4
A:Reference number: A26038; MUID:87115821; PMID:3027575
A:Accession: A26038
A:Molecule type: mRNA
A:Residues: 1-457 <LIT>
A:Cross-references: GB:X04836; NID:g50353; PIDN:CAA28539.1; PID:g50354
R:Gorman, S.D.; Tourville, B.; Parnes, J.R.
Proc. Natl. Acad. Sci. U.S.A. 84, 7644-7648, 1987
A:Title: Structure of the mouse gene encoding CD4 and an unusual transcript in brain.
A:Reference number: A39893; MUID:88041159; PMID:2823269
A:Accession: A39893
A:Molecule type: DNA
A:Residues: 1-25, 'E', 27-457 <GOR>
A:Cross-references: GB:M17080; GB:J03003; NID:g192515; PIDN:AAA37402.1; PID:g387124
R:Maddon, P.J.; Molineaux, S.M.; Maddon, D.E.; Zimmerman, K.A.; Godfrey, M.; Alt, F.W.;
Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159, 1987
A:Title: Structure and expression of the human and mouse T4 genes.
A:Reference number: A39955; MUID:88097446; PMID:3501122
A:Accession: A39955
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 25-457 <MAD>
A:Note: the cited GenBank accession number, J03564, is not in release 101.0
R:Parnes, J.R.; Hunkapiller, T.
Immunol. Rev. 100, 109-127, 1987
A:Title: L3T4 and the immunoglobulin gene superfamily: New relationships between the im
A:Reference number: I54564; MUID:88152875; PMID:3326818
A:Accession: I54564
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-457 <RES>
A:Cross-references: GB:M36850; NID:g198670; PIDN:AAA39401.1; PID:g198671
A:Accession: I69018
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 208-318 <RE2>
A:Cross-references: GB:M36851; NID:g198672; PIDN:AAA39402.1; PID:g554183
R:Classon, B.J.; Teagaratso, J.; Kirszbaum, L.; Maddox, J.; Mackay, C.R.; Brandon, M.;
Immunogenetics 23, 129-132, 1986
A:Title: The L3T4 antigen in mouse and the sheep equivalent are immunoglobulin-like.
A:Reference number: A47642; MUID:86166694; PMID:3082751
A:Accession: A47642
A:Molecule type: protein
A:Residues: 27-43 <CLA>
C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells
C:Genetics: 18/1, 74/1, 128/1, 207/1, 319/1, 386/1, 425/3, 448/2
A:Introns: 18/1, 74/1, 128/1, 207/1, 319/1, 386/1, 425/3, 448/2
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
C:Keywords: alternative initiators; duplication; glycoprotein; T-cell; transmembrane pr
F1-26/Domain: signal sequence #status predicted <SIG>
F:27-457/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>

[illegible]

A;Cross-references: GB:M59488
C;Genetics:
A;Gene: FlyBase:Slx1
A;Cross-references: FlyBase:FBgn0003659

Query Match 5.8%; Score 108.5; DB 2; Length 355;
Best Local Similarity 21.3%; Pred. No. 0.28;
Matches 67; Conservative 46; Mismatches 110; Indels 91; Gaps 13;

QY 5 LGTLLAATLAPFGAMGIVITGNHVSARID-----DDHIVIVAPREATIQ 51
DB 54 LGNVA-AHNISPGSNNINNTSNTNINFTIRONGVAALHYLOEQLOQPOQOQQ 111

QY 52 LQLFFPGORPHK-----YSGTVRVARSDITNQCQELSEERFENCNTHSSSVFVGC 105
DB 112 QOPLTWPPSPPPQOQSRQSHNGS-----SSILGNQLLAISNNNSFNNSNOSN-FTGN 165

QY 106 KVEYTFASNRLTG---PPHP-----FKLITRNPRPNDGMF-----YVIVR 145
DB 166 YNGSAFTNGAISGNFPNPNFTSSGNFTNNTSNTNFTNSGHSFASNLAGSSNFTNLSGS 225

QY 146 LDTKEPIDVFAIQLSVYOFANTATRGLYSK-----ASCR----- 181
DB 226 NNYTNSNGNFTSNAASSNFSNNAASSTNYKNCSSGVVGNSDPDSCSKHTTNTNNEHT 285

QY 182 -----TFGLPTVQLAYLRT-----EESRWQAVATEATTSAEATTPVTAT 227
DB 286 SPOHDFSPNMSTPEQLHQOTLKLOQLHLNNSFNN-----TTAAATTSAAASTSTTTAT 340

QY 228 SASELEAEHFTFPW 241
DB 341 -ASSTNSANVGFLW 353

RESULT 13
JN0133
WD-40 repeat regulatory protein tup1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: AER2 protein; flocculation suppressor protein SFL2; protein YCR084c
C;Species: Saccharomyces cerevisiae
C;Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text change 21-Jul-2000
C;Accession: JN0133; S19499; S13308; S11169; A36362; S13755
R;Zhang, M.; Rosenblum-Vos, L.S.; Lowry, C.V.; Boakye, K.A.; Zitomer, R.S.
Gene 97, 153-161, 1991
A;Title: A yeast protein with homology to the beta-subunit of G proteins is involved in
A;Reference number: JN0133; MUID:91153643; PMID:1900249
A;Accession: JN0133
A;Molecule type: DNA
A;Residues: 1-713 <ZHA>
A;Cross-references: GB:M35861; NID:g171037; PIDN:AAA34413.1; PID:g171038
R;Dusterhoft, A.; Erdmann, D.; Hegemann, J.; Philippesen, P.; Schweitzer, B.; Spiegelberg
submitted to the Protein Sequence Database, March 1992
A;Reference number: S19351
A;Accession: S19499
A;Molecule type: DNA
A;Residues: 1-713 <DUS>
A;Cross-references: EMBL:X59720; NID:g1907116; PID:e364571; PID:g1907221; MIPS:YCR084c
R;Williams, P.E.; Trumbly, R.J.
Mol. Cell. Biol. 10, 6500-6511, 1990
A;Title: Characterization of TUP1, a mediator of glucose repression in Saccharomyces cere
A;Reference number: A36362; MUID:91061758; PMID:2247069
A;Accession: S13308
A;Molecule type: DNA
A;Residues: 1-74, 'A', 76-99, 'Q', 101-684, 'P', 686-713 <WIL>
A;Cross-references: EMBL:M31733
R;Fujita, A.; Matsumoto, S.; Kuhara, S.; Misumi, Y.; Kobayashi, H.
Gene 89, 93-99, 1990
A;Title: Cloning of the yeast SFL2 gene: its disruption results in pleiotropic phenotype
A;Reference number: S11169; MUID:90323611; PMID:2197185
A;Accession: S11169
A;Molecule type: DNA
A;Residues: 45-713 <FUJ>
A;Cross-references: EMBL:X16365; NID:94459; PIDN:CAA34411.1; PID:g4460
C;Comment: This protein is responsible for maintaining repression of at least two distin

C;Genetics:
A;Gene: SGD:TUP1; ABR2; SFL2
A;Cross-references: SGD:S0000680; MIPS:YCR084c
A;Map position: 3R
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
C;Keywords: transcription regulation
F;181-198/Region: glutamine-rich
F;409-422/Region: threonine-rich
F;439-472/Domain: WD repeat homology <WD1>
F;522-556/Domain: WD repeat homology <WD3>
F;572-605/Domain: WD repeat homology <WD4>

Query Match 5.8%; Score 108; DB 2; Length 713;
Best Local Similarity 27.3%; Pred. No. 0.76;
Matches 62; Conservative 30; Mismatches 89; Indels 46; Gaps 11;

QY 60 QRPKPVGVTVRVARSDITNQCQELSEERFENCNTHSSSVFVGCVTYTF 115
DB 86 QRDHQIASLTVQOQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQO 144

QY 116 NRLTGPHPFKLITRNPRPNDGMFVIVRLDDEKPI-----D 154
DB 145 NNTTGSAPF--PVQASRPNLVGS-----QLPTTTLFVSSNAQQQLPQOQLQQO 197

QY 155 VFAIQLSVYOFANTATRGLYSKASCTFGLPTVQL-EAYLRTESWRNQA 208
DB 198 QPPQVQVAVLSNVAING---SPTSKETTLPSKAPESLTKETEPENNTSKINDT 254

QY 209 TEATTTSAEAT--TPTEVTATSAEAEHFTFPWLENGVDHYEPTP 253
DB 255 TTATTTTATETETKEEDATPAS-LHQDHYLVFYNQR-ANHSKPIP 299

RESULT 14
QOBE88
glycoprotein I precursor - human herpesvirus 2 (strain 333)
C;Species: human herpesvirus 2
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C;Accession: A05246
R;Hodgman, T.C.; Minson, A.C.
Virology 153, 1-11, 1986
A;Title: The herpes simplex virus type 2 equivalent of the herpes simplex virus type 1
A;Reference number: A05246; MUID:86291145; PMID:3016980
A;Accession: A05246
A;Molecule type: DNA
A;Residues: 1-372 <HOD>
A;Cross-references: EMBL:M14886; NID:g330311; PIDN:AAA45861.1; PID:g330313
C;Genetics:
A;Map position: 0.919-0.927
C;Superfamily: herpesvirus US7 protein
C;Keywords: glycoprotein; transmembrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-372/Product: glycoprotein I #status predicted <GPI>
F;255-282/Domain: transmembrane #status predicted <TM>
F;156,169,175,243/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.7%; Score 107; DB 1; Length 372;
Best Local Similarity 21.0%; Pred. No. 0.4;
Matches 74; Conservative 51; Mismatches 163; Indels 64; Gaps 13;

QY 3 SILGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQ-----QLPFF 56
DB 5 SLOQ-LAILGLWVC---ATGLVVRGPTVSL-VSDSLVDAGAVGPGFVEDLRVFGELHF 59

QY 57 MFGQRHPKPVGVTVRVARSDITNQCQELSEERFENCNTHSSSVFVGCVTYTF 116
DB 60 VQAQVPHNTNYDGIIEIIEFHYPLGNHCRPVHVHTLTACPRPAVFTLCRSTHAHS--- 116

QY 117 RLGTGPHHPF-----KLITRNPRNDGMFVIVRLDD-TKEPIDVFAIQLSV 166
DB 117 ----PAYPTLEGLARQPLLRVTRATRDYAGLVLRVWVGSATNASLFLVGLVALS---A 168

Search completed: March 8, 2004, 02:45:23
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 01:22:05 ; Search time 27 Seconds
(without alignments)
698.126 Million cell updates/sec

Title: US-09-993-777-7

Perfect score: 1863
Sequence: 1 MASLGTLLAATLAPFGA.....VELVAIVNPSALSPPDSIKM 362

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	162.5	8.7	354	1 VGLI_VZVD	P09258 varicella-z
2	155	8.3	353	1 VGLI_SVVD	Q04547 simian vari
3	153.5	8.2	424	1 VGLI_HSVB	P18553 equine herp
4	143.5	7.7	350	1 VGLI_PVR1	P07646 pseudorabie
5	140.5	7.5	380	1 VGLI_HSVB	Q08102 bovine herp
6	115	6.2	457	1 CD4_MOUSE	P06332 mus musculu
7	113.5	6.1	390	1 VGLI_HSV11	P06487 herpes simp
8	108	5.8	713	1 TUP1_YEAST	P16649 saccharomyc
9	107	5.7	372	1 VGLI_HSV23	P06764 herpes simp
10	99	5.3	372	1 VGLI_HSV2H	P13291 herpes simp
11	99	5.3	1036	1 P200_MYCPN	P75211 mycoplasma
12	98.5	5.3	1140	1 YD76_MYCPN	P75405 mycoplasma
13	98	5.3	503	1 CD44_RAT	P26051 rattus norv
14	98	5.3	507	1 YG46_YEAST	P53301 saccharomyc
15	97	5.2	1220	1 SLN1_YEAST	P39928 saccharomyc
16	95.5	5.1	797	1 VGLX_HSVB	P28968 equine herp
17	93.5	5.0	907	1 VGP3_EBV	P03200 Epstein-bar
18	93	5.0	1051	1 TF1A_MOUSE	Q64127 mus musculu
19	91	4.9	836	1 NOT3_YEAST	P06102 saccharomyc
20	91	4.9	3178	1 Y889_CABEL	Q09624 caenorhabdi
21	90.5	4.9	547	1 ICA3_HUMAN	P32942 homo sapien
22	90.5	4.9	785	1 CPTA_CABEL	Q17758 caenorhabdi
23	90.5	4.9	1429	1 YC31_SCHPO	Q14065 schizosacch
24	88.5	4.8	344	1 ULB6_HCMVA	P16833 human cytom
25	88	4.7	376	1 FOS_FUGRU	P53450 fugu rubrip
26	86.5	4.6	1037	1 YH1V_ECOLI	P37637 escherichia
27	86	4.6	430	1 CINA_MYCTU	Q07731 mycobacteri
28	86	4.6	636	1 VNR6_YEAST	P53882 saccharomyc
29	85.5	4.6	503	1 PODX_MOUSE	Q9R0M4 mus musculu
30	85	4.6	376	1 MID2_YEAST	P36027 saccharomyc
31	85	4.6	583	1 CH12_CANAL	P40953 candida alb
32	85	4.6	812	1 FGRI_XENLA	P22182 xenopus lae
33	85	4.6	1718	1 RRPO_SHVX	Q04575 shallot vir

ALIGNMENTS

RESULT 1

VGLI_VZVD	STANDARD;	PRT;	354 AA.
ID VGLI_VZVD			
AC P09258;			
DT 01-MAR-1989 (Rel. 10, Created)			
DT 01-MAR-1989 (Rel. 10, Last sequence update)			
DT 01-OCT-1996 (Rel. 34, Last annotation update)			
DE Glycoprotein I precursor (Glycoprotein IV) (GI) (GPV).			
GN 67.			
OS Varicella-zoster virus (strain Dumas) (VZV).			
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC Alphaherpesvirinae; Varicellovirus.			
OX NCBI_TaxID=10338;			
RN (1)			
RP SEQUENCE FROM N.A.			
RX MEDLINE=86306657; PubMed=3018124;			
RA Davison A.J., Scott J.E.;			
RT "The complete DNA sequence of varicella-zoster virus.;"			
RL J. Gen. Virol. 67:1759-1816(1986).			
RN (2)			
RP SEQUENCE FROM N.A.			
RX MEDLINE=84131932; PubMed=6321154;			
RA Davison A.J.;			
RT "DNA sequence of the US component of the varicella-zoster virus genome.;"			
RL EMOB J. 2:2203-2209(1983).			
RN (3)			
RP MUTAGENESIS, AND PHOSPHORYLATION SITE.			
RX MEDLINE=94267879; PubMed=8207795;			
RA Yao Z., Grose C.;			
RT "Unusual phosphorylation sequence in the gpI (gI) component of the varicella-zoster virus gpI-gpIV glycoprotein complex (VZV gE-gI complex).;"			
RL J. Virol. 68:4204-4211(1994).			
CC -!- SUBUNIT: HETERODIMER OF GPV AND GPV (GE).			
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.			
CC -!- SIMILARITY: TO HERPESVIRUSES GLYCOPROTEINS I, AND TO PRV GP63.			
CC -----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL; X04370; CAA27950.1; -			
DR EMBL; X00208; CAA25032.1; -			
DR F01; F27345; VGBB67.			
DR InterPro; IPR002874; Herpes_gI.			
DR InterPro; IPR007110; Ig-like.			
DR Pfam; PF01688; Herpes_gI; 1.			
DR Glycoprotein; Phosphorylation; Transmembrane; Signal.			
FT SIGNAL 1 17			
FT CHAIN 18 354			
FT DOMAIN 18 295			
FT GLYCOPROTEIN I.			
FT EXTRACELLULAR (POTENTIAL).			

O14763 homo sapien
P13463 agrobacteri
P08648 homo sapien
P22314 homo sapien
Q987B4 rattus norv
O88799 mus musculu
P14644 rattus norv
Q84YC2 xenopus lae
P53675 homo sapien
P33922 escherichia
P15379 mus musculu
P47033 saccharomyc

```

FT TRANSMEM 296 312 POTENTIAL.
FT DOMAIN 313 354 CYTOPLASMIC (POTENTIAL).
FT MOD RES 343 343 PHOSPHORYLATION.
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 T-A: NO EFFECT ON PHOSPHORYLATION.
FT MUTAGEN 341 342 EE-A: MODERATE DECREASE IN PHOSPHORYLATION.
FT MUTAGEN 343 343 S-A: MARKED DECREASE IN PHOSPHORYLATION.
FT MUTAGEN 344 344 P-A: MARKED DECREASE IN PHOSPHORYLATION.
FT MUTAGEN 345 345 P-A: DECREASE IN PHOSPHORYLATION.
SQ SEQUENCE 354 AA; 39364 MW; 55280608F52E666 CRC64;

Query Match 8.7%; Score 162.5; DB 1; Length 354;
Best Local Similarity 22.4%; Pred No. 4.8e-06;
Matches 72; Conservative 53; Mismatches 111; Indels 85; Gaps 15;

QY 23 IVITGNHVSARIDDDHIVIVAPRE---ATIQQLFFMPGQRP-HKPYSGTVVAFRSDI 78
D 21 LIFKGDHVSQVNSLSLILPMQNDNYTEIKQLVFIGQLPTGNTYSOTLELLY-ADT 79

QY 79 TNOCYQELSEBERFENCTHRSSVFGCKVT---EYTFASNRLTGPPHFKLTIRPREN 135
D 80 VAFCFRSGVIRYDGCPRITSAFISCKYKHSWHYGNSTDRISTEPDAGVMLKITKPGIN 139

QY 136 DSGMFYIVRLDTPKEPIDVFAQLSVY-----QFANTAATRGLYSKAS 179
D 140 DAGYVLLVRLDHSRS-TDGFILGVNVVITAGSHNHGVITSPSLONGVSTRALFQQA- 197

QY 180 CRTGGLPT-----VOLEAYLRTEES-WRNQAYVATEATTTSAEATTPPTVT 225
D 198 -RLCOLPATKSGSTSLFQWMLDRACKSLDNPWLH-EDVVTETKSVVKEG----- 248

QY 226 ATSASELEAEHFTFPLWENGVDHYEPTPAN---ENS-----NVTYRLGTMSPTLIGTVV 276
D 249 -----IEN---HYVPTDMSTLPEKSLNDPPENLLI-----IIPIVA 281

QY 277 AAVVSATIGLVIVISVTRNM 297
D 282 SVMILTAMVIVISVTRRRI 302

RESULT 2
VGLI SVVD STANDARD; PRY; 353 AA.
AC Q04547;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Glycoprotein I precursor (Membrane glycoprotein 1).
GN GI OR US3.
OS Simian varicella virus (strain DHV) (Cercopithecine herpesvirus 9).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=36348;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93212509; PubMed=9384754;
RA Fletcher T.M. III, Gray W.B.;
RT "DNA sequence and genetic organization of the unique short (US)
region of the simian varicella virus genome.";
RL Virology 193:762-773(1993).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEINS 1, TO VZV GP1V,
AND TO PRV GP63.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; L07067; AAA47888.1; -.
CC PIR; C46113; C46113.
DR InterPro; IPR002874; Herpes_g1.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF01688; Herpes_g1; 1. Signal.
KW Glycoprotein; Transmembrane; POTENTIAL.
FT SIGNAL 1 20
FT CHAIN 21 353 GLYCOPROTEIN I,
EXTRACELLULAR (POTENTIAL).
FT DOMAIN 21 274 POTENTIAL.
FT TRANSMEM 275 293
FT DOMAIN 294 353 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 353 AA; 40470 MW; E0314F7B5B0E4AA2 CRC64;

Query Match 8.3%; Score 155; DB 1; Length 353;
Best Local Similarity 20.2%; Pred No. 2e-05;
Matches 65; Conservative 66; Mismatches 151; Indels 40; Gaps 11;

QY 18 FG---AMGIVITGNHVSARIDDD---HIVIVAPREATIQQLFFMPGQRP-HKPYSGTVR 71
D 20 FGIQCAAIYRGVIVSYLVNSATSIFLKGNNNDASIRGLFIDGDFPVNTYNTVTE 79

QY 72 VAFRSDIINQCYQELSEBERFENCTHRSSVFGCKV-TEYTFASNRLTGPPHFKLTIR 130
D 80 L-LHVQNTLCLQPLRYWVYGECPRIITGAIACRVKRSWHYENATQUTDPNVEIIFKN 138

QY 131 NRPNDSCMFYIVRLDTPKEPIDVFAQLSVYQFANTAAARGLYSKASCRTFGLPTVOL 190
D 139 NTKVEDAGIYLLVQLDYT-SLPDIFVSLNVYPKQDTSNEDVNY-----FPPVYS 188

QY 191 EAY-LRTEESWRNQAYVATEATTTSAEATTPVTATTSASELEAEHFTFPLWENGVDHY 249
D 189 PSHILNTFKICHKFPVHNGMEQSILOHIVT-----SDVTETENLSQKDDLGST 238

QY 250 EPTPANENSVTVRL-----GTMSPTLIGTVVAAVVSATIGLVIVISI-VTRNVC 298
D 239 QKPRKNFNPDKVNVVTHETKTLMESSADVFMIAVPITASLLVILAIIVTVGVYRRS 298

QY 299 TPRKLDTVSQDDEERSOTRRE 320
D 299 SEKRKIYRPRKTKEQASTEKRE 320

RESULT 3
VGLI HSVB STANDARD; PRY; 424 AA.
ID VGLI HSVB
AC P18553;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glycoprotein I precursor.
GN GI OR 73.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1),
OS Equine herpesvirus type 1 (strain AB1) (EHV-1), and
OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520, 10328, 10330;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AB4P;
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1.";
```


Virology 189:304-316(1992).

[2]

RN
RN
RC
RC
RX
RA
RT
RT
RT
RT
RN

SEQUENCE FROM N.A.
STRAIN=AB1;
MEDLINE=91276272; PubMed=1647359;
Elton D.M., Halliburton I.W., Killington R.A., Meredith D.M.,
Bonass W.A.;
"sequence analysis of the 4.7-kb BamHI-EcoRI fragment of the equine
herpesvirus type-1 short unique region.";
Gene 101:203-208(1991).
[3]
SEQUENCE FROM N.A.
STRAIN=Kentucky D;
MEDLINE=91108393; PubMed=2177089;
Audonnet J.-C., Winslow J., Allen G., Paletti E.;
"Equine herpesvirus type 1 unique short fragment encodes
glycoproteins with homology to herpes simplex virus type 1 gD, gI and
gE";
J. Gen. Virol. 71:2969-2978(1990).
CC -! SUBCELLULAR LOCATION: Type I membrane protein.
CC -! SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEINS I, TO VZV GP1V,
AND TO PRV GP63.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; M86664; AAO02508.1; -;
EMBL; M36299; AAA66547.1; -;
PIR; C36646; VGEBE9.
InterPro: IPR002874; Herpes_gI.
Fram; PF01688; Herpes_gI; 1.
KW Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 424 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 23 319 POTENTIAL.
FT TRANSMEM 320 340 POTENTIAL.
FT DOMAIN 341 424 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 131 131 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
SEQUENCE 424 AA; 46392 MW; EE9BEFF7DA8A95806 CRC64;

Query Match 8.2%; Score 153.5; DB 1; Length 424;
Best Local Similarity 23.8%; Pred. No. 3.3e-05;
Matches 70; Conservative 41; Mismatches 126; Indels 57; Gaps 12;

QY 1 MASLLGTALLATLAPFGAMIVITGNHVSARIDDHHVIVAVPEPTIQQLFPMPG 59
DB 1 MAKLTGFSAAILLSMAICTAIYRGHEMSYLNASSEFAFYTPDQSILVLGHLLFDG 60

QY 60 QR-PHPYSGTVAFRSDITNQCYQEELSEERFENCNTSRSSSVFVC--KVETTFGSAN 116
DB QRLFTNYSLIEL-IHYNISVCYTIVIGTISYSCPRVANNAFRSCLHKTSKHVDYFR 119

QY 117 RLTGPPPHFKLTIRNPNDSGMFYVVRDLDTKEPIDVFATQLGVYOFANTAARGLYS 176
DB VNASVTENVNLTKOPTDSGAYLRVKLDIA-FADVFGVSFAFYDL----- 167

QY 177 KASCRTFGLPTVG----LEAYLTRTESWRWQAAYATEATTTS---AEATTTPTPVTAISA 229
DB KSXKVPDPMTTCTVBTPTSYSYSTPYDYDDVVTTETESTSTSQTQATSITQP-SATWG 226

QY 230 SELEAEHFTHPPWLNGVDHYEPYPANNSNVTVRLG-----TMSPTL 271

227 TOLITE-----IPTNE-----TWVIGCEALLCHWFQFSRTRVPFL 260

RESULT 4
VGLI PRIVI
ID VGLI PRIVI STANDARD; PRT; 350 AA.
AC P07646;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glycoprotein GP63 precursor
OS Pseudorabies virus (strain Rice) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10350;
RN [1] SEQUENCE FROM N.A.
RP MEDLINE=86308235; PubMed=3018284;
RA Petrovskis E.A., Timmins J.G., Post L.E.;
RT "Use of lambda gt11 to isolate genes for two pseudorabies virus glycoproteins with homology to herpes simplex virus and varicella-zoster virus glycoproteins.";
RT zoster virus glycoproteins.";
RL J. Virol. 60:185-193(1986).
CC -! SUBCELLULAR LOCATION: Type I membrane protein.
CC -! SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEINS I, AND TO VZV GP1V.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; M14336; AAC35204.1; -;
PIR; A29012; VGEB83.
InterPro: IPR002874; Herpes_gI.
Fram; PF01688; Herpes_gI; 1.
KW Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 350 GLYCOPROTEIN GP63.
FT DOMAIN 24 285 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 286 308 POTENTIAL.
FT DOMAIN 309 350 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 56 56 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
SEQUENCE 350 AA; 36773 MW; 66AF2229EC21BEDA CRC64;

Query Match 7.7%; Score 143.5; DB 1; Length 350;
Best Local Similarity 22.8%; Pred. NO. 0.00017;
Matches 82; Conservative 42; Mismatches 149; Indels 87; Gaps 17;

QY 8 LALLAATLAPFGM--GIIVTGNHV SARIDDHHVIVAVPREATIQLFPMPGORPKP 65
DB LLALAAATLAALTTPRVGVGLFRAGVSVHAGSAVLPGDPAFNLTIDGTLLEGPSFN- 73

QY 66 YSGTVRAFRSDITNQCYQEELSEERFENCNTSRSSSVFVC--KVETTFGSAN-----NRL 118
DB YSGRVEL-LRLDFPKKACTREYAASYDIQCPRHVEAFRCGLURKEPLARRASAARELL 132

QY 119 TGPPPHFKLTIRNPNDSGMFYVVRDLDTKEPIDVFATQLGVYOFANTAARGLYS 174
DB -----LFVERPAPPDAGSVLVVRVNGT---TDLPVLTALV-----PPGRPHHP 174

QY 175 --YSKASCRFTGLPTVQLEAVLTRTESWRWNQAYV--ATEATTTS---AEATTTPTPV 224
DB TPSSADECR-----PVV-----GSHWDLSRVDPADVFTTPPTTTPPP 221


```
QY 225 TATSAS-----ELEAHHFTFPLLENGVDHYEPTPANENSVTVRLGTMSPTLIGTVTA 277
Db 222 RGTGATPERSDEEED-----EGATTATPVPGLDANGTM---VLNASVSVRL 271
QY 278 AVVSATIG-----LVIVISIVTRNMCTPHRKLDTVSQDDERSQFRRESRX 323
Db 272 AAANATAGARGPKIAMVLGETIVVLLIFLGVAACARRCARGIAGTGRDGPGAARSTR 331
RESULT 5
VGLI_HSVBS STANDARD; PRT; 380 AA.
AC Q08102;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE Glycoprotein 1.
GN Gi.
OS Bovine herpesvirus type 1.2 (strain ST).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=45407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94167875; PubMed=8122370;
RA Leung-Tack P., Audonnet J.F., Riviere M.;
RT "The complete DNA sequence and the genetic organization of the short
unique region (US) of the bovine herpesvirus type 1 (ST strain).";
RL Virology 199;409-421(1994).
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEINS 1, TO VZV GP1V,
AND TO PRV GP63.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 223068; CAA0605.1; -.
DR F01; S35785; S35785.
DR InterPro; IPR002874; Herpes_g1.
DR Pfam; PF01688; Herpes_g1; 1.
KW Glycoprotein.
FT CARSDYD 67 67 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 380 AA; 39910 MW; ABE1FB9B430D2BD CRC64;
Query Match
Best Local Similarity 7.5%; Score 140.5; DB 1; Length 380;
Matches 68; Conservative 33; Mismatches 125; Indels 43; Gaps 9;
QY 1 MASLIGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPPEATQL--QLFFMP 58
Db 1 MRCILLMMVLAARAP--ARSLYRGEAVGLRADGPFVAFVHP-TDATALRGLRFLFE 57
QY 59 GQRP-HKYSCTGVVARSDITNOCYQELSEEPENCETHSVFFGCKVTEFFPSNR 117
Db 58 HQLPAGRYNGTVEL-LRYHAAGDFVLMQTTAFASPRVANNAPRSLHADTRPARSER 116
QY 118 LTGP--PHPFKLTIRNRPDNDGFMFYIVRL-----DDTKEPIDVFAQLSVQFA---- 166
Db 117 RASAAVENHVLFSIARPPIDSGLYFLRVGIYGTAGSERRRDVFPLAFVHSFGPGDP 176
QY 167 -----NTAATRGLYSKASCTFGI-----PTVQLEAVLRTEESWR 201
Db 177 EAAARTPAPRQSRPAASGLTSSASLYDALARSPOAPPRPAPPAARAGPRRPERVDE 236
QY 202 NQAVVATEATTSEATPTTFVATAS 230
Db 237 TTEVEAATRASAFALTPPAGPTASPA 265
```

RESULT 6

```
CD4_MOUSE
ID CD4_MOUSE STANDARD; PRT; 457 AA.
AC P06332;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3) (T-cell differentiation antigen L3T4).
DE CD4.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87018845; PubMed=3094146;
RA Tourville B., Gorman S.D., Field E.H., Hunkapiller T., Parnes J.R.;
RT "Isolation and sequence of L3T4 complementary DNA clones: expression
in T cells and brain.";
RL Science 234;610-614(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87115821; PubMed=3027575;
RA Littman D.R., Gattner S.N.;
RT "Unusual intron in the immunoglobulin domain of the newly isolated
murine CD4 (L3T4) gene.";
RL Nature 325;453-455(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=88152875; PubMed=3326818;
RA Parnes J.R., Hunkapiller T.;
RT "L3T4 and the immunoglobulin gene superfamily: new relationships
between the immune system and the nervous system.";
RL Immunol. Rev. 100;109-127(1987).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=88041159; PubMed=2823269;
RA Gorman S.D., Tourville B., Parnes J.R.;
RT "Structure of the mouse gene encoding CD4 and an unusual transcript
in brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 84;7644-7648(1987).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=98112780; PubMed=9445485;
RA Ansari-Lari M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,
Lu J., Gorrell J.H., Chinault A.C., Belmont J.W., Miller W.,
Gibbs R.A.;
RT "Comparative sequence analysis of a gene-rich cluster at human
chromosome 12p13 and its syntenic region in mouse chromosome 6.";
RL Genome Res. 8;29-40(1998).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Uden T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Paney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE OF 27-43.
 RX MEDLINE=86166694; PubMed=3082751;
 RA Claesson B.J., Tsegaratos J., Kirszbaum L., Maddox J., McKay C.R.,
 RA Brandon M., McKenzie I.P.C., Walker I.D.;
 RT "The L374 antigen in mouse and the sheep equivalent are
 immunoglobulin-like.";
 RL Immunogenetics 23:129-132(1986).
 RN [8]
 RP DISULFIDE BONDS.
 RX MEDLINE=86233454; PubMed=3086886;
 RA Claesson B.J., Tsegaratos J., McKenzie I.P.C., Walker I.D.;
 RT "Partial primary structure of the T4 antigens of mouse and sheep:
 assignment of intrachain disulfide bonds.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:4499-4503(1986).
 CC -!- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 receptor interaction. May regulate T-cell activation.
 CC -!- SUBUNIT: Associates with p56-lck (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P06332-1; Sequence=Displayed;
 CC Name=2; Synonyms=Brain-specific;
 CC IsoId=P06332-2; Sequence=VSP 002489;
 CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 DR EMBL; M36850; AAA39401.1; -
 DR EMBL; M13816; AAA37267.1; -
 DR EMBL; M13816; AAA37267.1; -
 DR EMBL; X04836; CAA28539.1; -
 DR EMBL; M36851; AAA39402.1; -
 DR EMBL; M17080; AAA37403.1; -
 DR EMBL; M17078; AAA37403.1; JOINED.
 DR EMBL; M17079; AAA37403.1; JOINED.
 DR EMBL; AC002397; AAC36010.1; -
 DR EMBL; BC039137; AAC39137.1; -
 DR PIR; A02110; RWMST4.
 DR HSSP; P01730; 1WBR.
 DR MGD; MGI:88335; Cd4.
 DR GO; GO:0042101; C-T-cell receptor complex; ISS.
 DR GO; GO:0015026; F:coreceptor activity; ISS.
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . . ISS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0045058; P:T-cell selection; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ISS.
 DR InterPro; IPR000973; CD4_TCRG.
 DR InterPro; IPR007110; Ig-like.
 DR Pfam; PF00047; Ig_2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00406; IGV_1.
 DR PROSITE; PS00835; Ig LIKE; 1.
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate;
 KW Alternative splicing.
 FT SIGNAL 1 26
 FT CHAIN 27 457 T-CELL SURFACE GLYCOPROTEIN CD4.
 FT DOMAIN 27 394 EXTRACELLULAR (POTENTIAL).

FT	TRANSMEM	395	417	POTENTIAL.
FT	DOMAIN	418	457	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	27	128	IG-LIKE V-TYPE.
FT	DOMAIN	129	207	IG-LIKE C2-TYPE 1.
FT	DOMAIN	208	317	IG-LIKE C2-TYPE 2.
FT	DOMAIN	318	374	IG-LIKE C2-TYPE 3.
FT	CARBOHYD	187	187	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	298	298	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	323	323	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	392	392	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	42	112	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	159	188	
FT	DISULFID	328	370	
FT	LIPID	418	418	
FT	LIPID	421	421	
FT	VARSPLIC	1	240	
FT	SEQUENCE	457 AA;	51296 MW;	1B1DAV527CB00F33 CRC64;
Query Match		6.2%;	Score 115;	DB 1; Length 457;
Best Local Similarity		20.6%;	Pred. No. 0.055;	
Matches		76;	Conservative	54; Mismatches 133; Indels 106; Gaps 16;
QY	19	GAMGIVITGNHVSAIRDDDDHIVIVAPRPETATIQOLF---	FMPGQPHKPYSGTVVAFR	75
Db	93	GSFPLII--NKLKMEDSQTYICELENRKE-EVELWVKVTFSPGTSLLQCSLTLLDSN	149	
QY	76	SDIYNQCYQELSEERFENCETHRSSVYVCKKTEYTFESANRLTGPFPFKLTIRPNRPN	135	
Db	150	SKVSN-----PLTE-----CHKKGKVVSGKV-----LSMSNLRVQ	181	
QY	136	DSGMFYVIVRLDDTKPEIDVFAIQLSVYOFANTAATRGLYSKASCRTFGLPTVQLEAYLR	195	
Db	182	DSDFNCTVTLQDKK---NWFGTSLVLGFSQTAIT-AYKSGESAEFSPLNFAB---	233	
QY	196	TEESWRN--NOAVYATEATTSABEATTPVATSASELEAEHPTFPWLENGVDHVEPTP	253	
Db	234	-ENGWGLMWKA-----EKDSFFQFWISFSIKKEVSU	265	
QY	254	ANENSNVTVRLGTMSPTLIGTVAAVVSATIGLVIVISIVTRNMCTPHKLDIVSQDDEE	313	
Db	266	QKSTYDKLQIKETPLTLTKIPQVSLQFAGSGL-----TLTLDKGLHVEVNLVVMKVAQ	321	
QY	314	RSQTRRESRKFGPMVACEIN-----KGADQD---SELVELVAIVNPSA-----	353	
Db	322	LNNT-----LTCEVMGPTS PKMRLTLKQENGEARVSEBQKVVQVVAETGLWQCL	371	
QY	354	LSSPDSIKM	362	
Db	372	LSBGDKVKM	380	

RESULT 7
 VGLI_HSV11
 ID VGLI_HSV11 STANDARD; PRT; 390 AA.
 AC P06487;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Glycoprotein I.
 GN GI OR US7.
 OS Herpes simplex virus (type 1 / strain 17).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 CC NCBI_TaxID=10299;
 RN [1]
 RX SEQUENCE FROM N.A.
 RA MEDLINE=85160822; PubMed=2984429;
 RA McGeoch D.J., Dolan A., Donald S., Rixon F.J.;
 RT "Sequence determination and genetic content of the short unique
 region in the genome of herpes simplex virus type 1.";
 RL J. Mol. Biol. 181:1-13(1985).
 CC -!- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND

DR InterPro: IPR002874; Herpes_g1.
DR Pfam: PF01688; Herpes_g1; 1.
KW Glycoprotein.
FT CARBOHYD 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 372 AA; 39558 MW; CODA22CAB16B8D7 CRC64;

Query Match 5.3%; Score 99; DB 1; Length 372;
Best Local Similarity 21.0%; Pred. No. 0.88;
Matches 74; Conservative 49; Mismatches 165; Indels 64; Gaps 13;

QY 3 SILGTLALLAATLAPGANGIVITGNHVSARDDDDHIVIVAPREPTQL-----QLFF 56
Db 5 SLQG-LAILGLWVC---ATGLVVRGPTVSL-VSDSLVDAGAVGPGFVEEDLRVFGELHF 59

QY 57 MFCQRPKPYSGTVRVAFPSDITNOCYQELSEERFENCNTHRSSVFGCKTEYTFSSAN 116
Db 60 VQAQVPHYNYDGIIELFHYPLGNHCPRVHVVTITACPRRPAVAFILCRSTHHAHS--- 116

QY 117 RLTGPPHPF-----KLITRNPRNDPMGYVIVRLDD-TKSPIDVFAIQLSVYQFA 166
Db 117 ----PAYPTLELGLARQPLLRVTATRDYAGLYLVLRVVGVSATNASRFLGVALS---A 168

QY 167 NTAATRGYSKASCTFGLPTVQLEAYLRTESNRNQAYVATEATTTSAEATTPPTVA 226
Db 169 NGTFVYNGSDYSCDPAQLP-----FSAPRLGPSVVTYPGASRTPPT 212

QY 227 TSASELEAHFTFPLWLVGVDHYEPTPANENSNTVRLGTMSPTLIGTVAAVVSATIGL 286
Db 213 TTPSSFRDPTAPG-----DTGTPAPASGETAPPNSTRSASESHRLTLVAQVQIAIPA 267

QY 287 VIVISIVTRN-MCTPHRKLDVTSQDDEERSQTRRSRKEFG-MVACEINKGA 336
Db 268 SIATFVLGSCICFIHRC-----QRRYRPRGQIINPGVSCAVNEAA 310

RESULT 11
ID P200_MYCPN STANDARD; PRT; 1036 AA.
AC P75211; Q50346;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein P200.
GN P200 OR MPN367 OR MP275.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=96257197; PubMed=8675035;
RA Proft T., Hilbert H., Plagens H., Herrmann R.;
RT "The P200 protein of Mycoplasma pneumoniae shows common features with
RT the cytochrome-associated proteins HMW1 and HMW3."
RL Gene 171:79-82(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Hammelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [3]
RP SEQUENCE OF 641-678 FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=95075318; PubMed=7984111;
RA Proft T., Herrmann R.;
RT "Identification and characterization of hitherto unknown Mycoplasma

pneumoniae proteins.";
RL Mol. Microbiol. 13:337-348(1994).
CC -!- FUNCTION: PROTEIN CYTOSKELETON-ASSOCIATED WHICH COULD BE AN
CC ACCESSORY STRUCTURAL COMPONENT IN CYTADHERENCE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to licenses@isb-sib.ch).
CC
DR EMBL; U25989; AAC99815.1; -
DR EMBL; AE00027; AAB95923.1; -
DR EMBL; Z32646; CAA83569.1; -
DR PIR; S73601; S73601.
KW Cytoadherence; Structural protein; Repeat; Complete proteome.
FT DOMAIN 277 280 POLY-THR.
FT DOMAIN 300 845 PRO-RICH.
FT DOMAIN 357 360 POLY-THR.
FT DOMAIN 401 404 POLY-ALA.
FT DOMAIN 718 781 3 X 6 AA REPEAT OF E-P-E-P-N-F.
FT REPEAT 718 723 1.
FT REPEAT 738 743 2.
FT REPEAT 776 781 3.
FT CONFLICT 641 641 A -> P (IN REF. 3).
SQ SEQUENCE 1036 AA; 116915 MW; DESAEAB6DD95B29 CRC64;

Query Match 5.3%; Score 99; DB 1; Length 1036;
Best Local Similarity 23.9%; Pred. No. 3.3;
Matches 57; Conservative 29; Mismatches 90; Indels 62; Gaps 10;

QY 127 LTRNPRNDPMGYVIVRLDDTKSPIDVFAIQLSVYQFAN---TAATRGYSKASCTF 183
Db 446 WIDNGQPQAGFYVVDFTLTST-APLTVAEIEIQEQLVNEFVTTTSTRTTTFASTPVF 504

QY 184 G---LPTVQLEAYLRTESNRNQAYVATEATTTSAEATTP---TPVTATSAEAEAH 236
Db 505 EPVVIPTVSEEQLENE-----FVSTVSATSNEPNAVSTPVET----- 546

QY 237 FTTFWLVGVDHYEPTPANENSNTVRLGTMSPTLIGTVV-----AAVVSATIGL 286
Db 547 -----VELTE-TPVSLPELETQLETPAVVTEITVTEKAVEPLVAVVEAPLAV 596

QY 287 VIVISIVTRNMCTPHRKLDVTSQDDEERSQTRRSRKEFG--PMVACEINKGADDSSEL 342
Db 597 EPIVETST-----TLAETVEEAQVEQSTAVAVEPALETESKATSEAQAE 643

RESULT 12
ID YD76_MYCPN STANDARD; PRT; 1140 AA.
AC P75405;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN376 (A19_orf1140).
GN MPN376 OR MP460.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Hammelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; STRONG, TO M.PNEUMONIAE
CC MPN375.
CC

Cell 65:13-24 (1991).

[2]

SEQUENCE FROM N.A. (ISOFORM 1).

Stevens J.W., Midura R.J.;
Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

-|- FUNCTION: Main cell surface receptor for hyaluronate. Adhesion to
mucosal high endothelial venule and to types I and VI collagen.
Probably involved in matrix adhesion, lymphocyte activation and
lymph node homing.

-|- SUBCELLULAR LOCATION: Type I membrane protein.

-|- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Comment=Additional isoforms seem to exist;
Names=2; Synonyms=Long, Meta-1;
IsoId=p26051-1; Sequence=Displayed;
Name=1; Synonyms=Short;
IsoId=p26051-2; Sequence=VSP_005330;
-|- PFM: N-glycosylated (By similarity).
-|- PFM: O-glycosylated; Contains chondroitin sulfate glycans which
can be more or less sulfated (By similarity).
-|- PFM: Phosphorylated; activation of PKC results in the
dephosphorylation of Ser-467 (constitutive phosphorylation site),
and the phosphorylation of Ser-433 (By similarity).
-|- SIMILARITY: Contains 1 link domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M61875; AAA3532.1; -;
EMBL; M61874; AAA3534.1; -;
EMBL; U52179; AAA97915.1; -;
EMBL; U46957; AAA92920.1; -;
PIR; B38745; B38745.
HSSP; P98066; ITSG.
InterPro; IPR001231; CD44 antigen.
InterPro; IPR000538; Link.
Pfam; PF00193; Xlink; 1.
PRINTS; PR00658; CD44.
PRINTS; PR01265; LINKMODULE.
PRODOM; PD000919; Link; 1.
SMART; SM00445; LINK; 1.
PROSITE; PS01241; LINK; 1.
Cell adhesion; Transmembrane; Glycoprotein; Phosphorylation; Receptor;
Proteoglycan; Signal; Alternative splicing;
Pyroglutamate carboxylic acid.
SIGNAL 1 21
CHAIN 22 503 CD44 ANTIGEN.
DOMAIN 22 410 EXTRACELLULAR (POTENTIAL).
TRANSMEM 411 431 POTENTIAL.
DOMAIN 432 503 CYTOPLASMIC (POTENTIAL).
DOMAIN 51 123 LINK.
DOMAIN 154 162 ARG/LYS-RICH (BASIC).
DOMAIN 228 410 STEM.
MOD_RES 22 22 PYROLIDONE CARBOXYLIC ACID (PROBABLE).
DISULFID 56 122 BY SIMILARITY.
DISULFID 80 100 BY SIMILARITY.
MOD_RES 433 433 PHOSPHORYLATION (BY PKC) (PARTIAL) (BY
SIMILARITY).
MOD_RES 467 467 PHOSPHORYLATION (PARTIAL) (BY
SIMILARITY).
CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPLIC 224 385 IATPPWSARTKQKQERTQWNPISHNPEVLLQTTTRMTDID
RNTSAGHNTWQEQPOPPFNHEYODEETPHATITWADP
NSTTEAATQKSKWENWQGNPPTPSDSHVTRGTASA
HNHPQRMWTTQSDQVSWTDFDPIHPMGQGHQTESK
-> SDGSSMDPRGGDFIVHGSELA (in isoform
1).
FT ID: VSP 005330.
FT CONFLICT 74 74 R -> S (IN REF. 2).
FT SEQUENCE 503 AA; 55945 MW; FB489D009BD4EE22 CRC64;
Query Match 5.3%; Score 98; DB 1; Length 503;
Best Local Similarity 22.2%; Pred. No. 1.6;
Matches 47; Conservative 30; Mismatches 79; Indels 56; Gaps 10;
QY 178 ASCRTGLPVOLEAVLRTEESW--RWQO-----AYVATEATTTSAEATTPVT 225
DB 297 ATSTTADPNSTTTEAATQKKEFWENWQGNPPTPSDSHV-TEGTTASAHNHPQRM 355
QY 226 ATGASELEABHFTFPWLENGVDHYEP-----TPANENSNVTVRLGTWSP 269
DB 356 TTQSQE-----DVSW-----TDFFDPIHPMGQGHQTESKHSNGQDSGVTTTSGPARR 405
QY 270 TLIG---VTVAVVSATIGLVIVISIVTRNMCTPHRKLPTVQDDRESCQTRRGRKFGP 326
DB 406 PQIPEWLIILASLLALILAVCIANVRRCQKKKL--VINGNGTVEDRKPS----- 458
QY 327 MVACEINKGADQSELVELVAIVNPSALSPD 358
DB 459 ----ELNGEASKQEMVHL---VNKEPTETPD 483

RESULT 14
YQ46 YEAST
ID YQ46 YEAST STANDARD; PRT; 507 AA.
AC P53301;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 52.8 kDa protein in BUB1-HiP1 intergenic region.
GN YGR189C OR G7553.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=97279231; PubMed=9133739;
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
Nombela C.;
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
of Saccharomyces cerevisiae chromosome VII.";
RL Yeast 13:357-363(1997).
CC -!- SIMILARITY: SOME, TO YEAST UTR2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 272974; CAA97215.1; -;
DR EMBL; X99074; CAA67525.1; -;
DR PIR; S64507; S64507.
DR HSSP; P23904; LAJK.
DR GernOnline; 141501; -;
DR SGD; S0003421; CRH1.
DR GO; GO:0009277; C:cell wall (sensu Fungi); IDA.
DR GO; GO:0000131; C:incipient bud site; IDA.
DR InterPro; IPR008985; ConA like lec.gl.
DR InterPro; IPR000757; Glyco_hydro_16.

DR Pfam; PF00722; Glyco_hydro_16; 1.
KW Hypothetical protein.
FT DOMAIN 63 66 POLY-SER.
FT DOMAIN 301 310 POLY-SER.
FT DOMAIN 345 357 POLY-SER.
FT DOMAIN 387 391 POLY-SER.
FT DOMAIN 467 470 POLY-SER.
SQ SEQUENCE 507 AA; 52757 MW; 7D7B61F57ABA942C CRC64;
Query Match 5.3%; Score 98; DB 1; Length 507;
Best Local Similarity 21.7%; Pred. No. 1.6;
Matches 54; Conservative 46; Mismatches 95; Indels 54; Gaps 9;
QY 136 DSGMFFYVRLDDTKEPIDVFAIQLSVYQFANTAAATGLYSKASCTFGTGLPTVOLEAYLR 195
DB 278 DGGSIY--GRYDQAEQFAVLANGGSISSSTSSSTVSSSASTVSSSVSTVSSSASST 335
QY 196 TEESWRNQAVATEATTSAEATTPVTATATSELEAEHFTFWLENGVDHYEPTPAN 255
DB 336 VSSS---VSSVSSSVSSSSSTSPSSSTATSKTLASSVT---TSSSISSFE--KQSS 388
QY 256 ENSNVTVRLGTWSPTLIGVT--VAAVVSATIGLVIVISIVTRNMCTPHRKLDTVQSD--- 310
DB 389 SSSKKTVAASSTSSSIISSTKTPATVST-----TRSTVAPTQSSVSDSPVQ 438
QY 311 -----DEERQTRRESRKFGPMV-----ACEINKGADQD-----S 340
DB 439 DKGGVATSSNDVTSSTTQISSKYTSTIQSSSEASTNSTVQISNGADLAQSLPREGKIFS 498
QY 341 ELVELVAIV 349
DB 499 VLVALALL 507

RESULT 15
SLN1 YEAST
ID SLN1 YEAST STANDARD; PRT; 1220 AA.
AC P39328;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Osmolality two-component system protein SLN1 (EC 2.7.3.-).
GN SLN1 OR YPD2 OR YII47C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / YPH1;
RX MEDLINE=94024010; PubMed=8211183;
RA Ota I.M., Varshavsky A.;
RT "A yeast protein similar to bacterial two-component regulators.";
RL Science 262:566-569(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Hornslett T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87(1997).
RN [3]
RP MUTAGENESIS OF HIS-576 AND ASP-1144.
RX MEDLINE=94239498; PubMed=8183345;
RA Maeda T., Wurgler-Murphy S.M., Saito H.;
RT "A two-component system that regulates an osmosensing MAP kinase
RT cascade in yeast.";
RL Nature 369:242-245(1994).
CC -!- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM

CC SLN1/SSK1 ACTIVATED BY CHANGES IN THE OSMOLARITY OF THE
CC EXTRACELLULAR ENVIRONMENT. THIS SYSTEM CONTROLS THE SSK2/SSK22->
CC PB22-HOG1 PATHWAY. INACTIVE SLN1 ALLOWS THE UNPHOSPHORYLATED SSK1
CC PROTEIN TO ACTIVATE SSK2 AND SSK22, TWO MAPKKS THAT FURTHER
CC STIMULATE THE PB22-HOG1 MAPK CASCADE. IN LOW OSMOLARITY MEDIA,
CC THE ACTIVATED SLN1 HISTIDINE KINASE REPRESENTS THE ACTIVATION OF THE
CC PB22-HOG1 KINASE CASCADE THROUGH PHOSPHORYLATION OF SSK1.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- PTM: Activation probably requires a transfer of a phosphate group
CC between a His in the transmitter domain and an Asp of the receiver
CC domain.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z38059; CAA86131.1; -;
CC EMBL; U01835; AAC48912.1; -;
CC FIR; S48387; S48387.
CC HSP; P06143; IUDR.
CC GermOnline; 139682; -;
CC SGD; S0001409; SLN1.
CC GO; GO:0004673; F:protein-histidine kinase activity; IDA.
CC GO; GO:0007234; P:osmosensory signaling pathway via two-compo. . .; IDA.
CC GO; GO:0042542; P:response to hydrogen peroxide; IMP.
CC InterPro; IPR003594; ATPbind ATPase.
CC InterPro; IPR004358; Bact sens pr C.
CC InterPro; IPR003661; His KinA_N.
CC InterPro; IPR005467; His_kinase.
CC InterPro; IPR001789; Response_reg.
CC Pfam; PF02518; HATPase_c; 1.
CC Pfam; PF00512; Hiska; 1.
CC PRINTS; PR00344; ECTRLENSOR.
CC ProDom; PD000039; Response_reg; 1.
CC SMART; SM00387; HATPase_c; 1.
CC SMART; SM00388; Hiska; 1.
CC SMART; SM00448; REC; 1.
CC PROSITE; PS50109; HIS_KIN; 1.
CC PROSITE; PS50110; RESPONSE_REGULATORY; 1.
CC Sensory transduction; Transferrase; Kinase; Phosphorylation;
CC Transmembrane.
CC KW DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 23 46 POTENTIAL.
CC DOMAIN 47 333 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 334 354 POTENTIAL.
CC DOMAIN 355 1220 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 573 928 HISTIDINE KINASE.
CC DOMAIN 1089 1210 RESPONSE REGULATORY.
CC MOD_RES 576 576 PHOSPHORYLATION (AUTO-) (PROBABLE).
CC MOD_RES 1144 1144 PHOSPHORYLATION (PROBABLE).
CC CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC MUTAGEN 891 891 H->Q: INACTIVE.
CC MUTAGEN 576 576 H->D: SLOW GROWTH, SLN1-1 MUTANT.
CC MUTAGEN 1144 1144 D->N: INACTIVE.
CC SEQUENCE 1220 AA; 134434 MW; 45FFE24A8165486B CRC64;

Query Match 5.2%; Score 97; DB 1; Length 1220;
Best Local Similarity 22.7%; Pred. No. 5.9;
Matches 58; Conservative 43; Mismatches 104; Indels 50; Gaps 11;
QY 5 LGTALLAATLAFGANGIVITGNHVSARIDDDHHIVAPRPEATLQLQFFPFGQRPKH 64

Db 32 LGSLLIILAVT-----TGVFTSNYNLR--SDRLVIAAQLKSSQIDQTLNVL-----YQ 79
QY 65 PYSGTVRVAFRSDITNQCYQELSEERFENCTH-----RSSSVFVGCKVTEVTF-----SA 114
Db 80 AYLASRDALQSSLTSYVAGNKSADNWDSLSVIOKFLSSSLFYVAKYDSSFNALVA 139
QY 115 SNELTGPPHP-----FKLTIRNPRPNDGMYFVIVRLDDTKERIDVFAIQLSVYQFAN 167
Db 140 TNNGTGDLIPEDVLDLSLPLSTDTPLP--SSLETIGITDPVLNSTD-YLMSMSLPIFAN 196
QY 168 TAATRGLYSKASCRTEGLPTVQLEAYLRTESWRNMQAVYATATTSABATTPVPTAT 227
Db 197 PSII-----LTDSRVGYITIIIMSA-----EGLKSVFNDTIALEHSTIAISAV 240
QY 228 SASELEAE--HFTFP 240
Db 241 YNSQKASGYHFVFP 255
Search completed: March 8, 2004, 02:42:18
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2004, 02:30:36 ; Search time 125 Seconds
(Without alignments)
913.741 Million cell updates/sec

Title: US-09-993-777-7

Perfect score: 1863

Sequence: 1 MASLIGTLALLAATLAPFGA.....VELVAIVNPSSALSSPDSIKM 362

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1863	100.0	362	Q67645	Q67645 gallid herp
2	164	8.8	420	Q39309	Q39309 equine herp
3	162.5	8.7	354	Q98VM1	Q98VM1 human herpe
4	158	8.5	420	Q55525	Q55525 equine herp
5	156	8.4	384	Q29357	Q29357 feline herp
6	154	8.3	384	Q66931	Q66931 feline herp
7	152	8.2	355	Q9E6L5	Q9E6L5 turkey herp
8	151	8.1	355	Q8YL08	Q8YL08 turkey herp
9	149	8.0	355	Q69287	Q69287 turkey herp
10	148	7.9	364	Q9YPA1	Q9YPA1 canine herp
11	148	7.9	364	Q41525	Q41525 canine herp
12	147	7.9	259	Q91335	Q91335 canine herp
13	145	7.8	370	Q86789	Q86789 feline herp
14	144	7.7	355	Q8JLU7	Q8JLU7 turkey herp
15	140.5	7.5	369	Q9IE39	Q9IE39 phocid herp
16	139.5	7.5	382	Q39505	Q39505 bovine herp

17	139	7.5	355	12	Q8JLU9	Q8JLU9 turkey herp
18	130.5	7.0	366	12	Q99F67	Q99F67 suid herpes
19	125	6.7	356	12	Q88524	Q88524 turkey herp
20	124.5	6.7	883	12	Q36401	Q36401 alcelaphine
21	123	6.6	353	12	Q67638	Q67638 gallid herp
22	115	6.2	433	11	Q55054	Q55054 mus musculu
23	115	6.2	457	11	Q61396	Q61396 mus musculu
24	112	6.0	387	4	Q86XK7	Q86XK7 homo sapien
25	112	6.0	582	5	Q8IR74	Q8IR74 drosophila
26	108.5	5.8	319	5	Q9W3S7	Q9W3S7 drosophila
27	108.5	5.8	475	16	Q8FQU1	Q8FQU1 corynebacte
28	108	5.8	381	5	Q8IR75	Q8IR75 drosophila
29	106.5	5.7	235	12	Q9DY19	Q9DY19 bovine herp
30	103	5.5	982	5	Q9VIG1	Q9VIG1 drosophila
31	102.5	5.5	378	4	Q96H15	Q96H15 homo sapien
32	102.5	5.5	4498	5	Q9W2Z3	Q9W2Z3 drosophila
33	100.5	5.4	2114	5	Q8IQ18	Q8IQ18 drosophila
34	99	5.3	279	5	Q9U474	Q9U474 caenorhabdi
35	99	5.3	1011	5	Q9NHZ9	Q9NHZ9 helicoverpa
36	98	5.3	264	11	Q80XS5	Q80XS5 mus musculu
37	98	5.3	862	11	Q8K0T1	Q8K0T1 mus musculu
38	98	5.3	1262	11	Q80T73	Q80T73 mus musculu
39	97.5	5.2	1711	5	Q9W1X4	Q9W1X4 drosophila
40	96.5	5.2	814	13	Q91897	Q91897 xenopus lae
41	96	5.2	877	4	Q9H3Q6	Q9H3Q6 homo sapien
42	96	5.2	878	4	Q9H3Q7	Q9H3Q7 homo sapien
43	96	5.2	957	4	O14651	O14651 homo sapien
44	96	5.2	1217	4	Q9UKW9	Q9UKW9 homo sapien
45	95.5	5.1	332	12	Q8V0L9	Q8V0L9 equine herp

ALIGNMENTS

RESULT 1

Q67645 PRELIMINARY; PRT; 362 AA.
 ID Q67645
 AC Q67645;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glycoprotein I.
 OS Gallid herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
 OX NCBI_TaxID=10386;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA challenge strain;
 RX MEDLINE=97033380; PubMed=8879127;
 RA Wild M.A., Cook S., Cochran M.;
 RT "A genomic map of infectious laryngotracheitis virus and the sequence
 and organization of genes present in the unique short and flanking
 regions."
 RL Virus Genes 12:107-116(1996).
 DR EMBL; U28832; AAC55101.1; .
 DR InterPro; IPR002874; Herpes_gI.
 DR InterPro; IPR007110; Ig-like.
 DR Pfam; PF01688; Herpes_gI; 1.
 SQ SEQUENCE 362 AA; 39750 MW; F530C1AAF7CC6BA5 CRC64;

Query Match 100.0%; Score 1863; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 1.9e-162;
 Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASLIGTLALLAATLAPFGAGVITGNHVSARIDDDHIVIVAPRPETATLQQLFFMEQG	60
Db	1	MASLIGTLALLAATLAPFGAGVITGNHVSARIDDDHIVIVAPRPETATLQQLFFMEQG	60
QY	61	RPHKPYSGTVRVAFPSDITNQCYOELSEERFENCTHRSSSVFVCKCTXTFFSASNLITG	120
Db	61	RPHKPYSGTVRVAFPSDITNQCYOELSEERFENCTHRSSSVFVCKCTXTFFSASNLITG	120

```
QY 121 PPHPFKLTIRNPRNDGMYFVIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASC 180
Db 121 PPHPFKLTIRNPRNDGMYFVIVRLDDTKEPIDVFAIQLSVYQFANTAATRGLYSKASC 180
QY 181 RTFGLPTVQLAYLRTESWRNQAYVATEATTTSAEATTPPTVATTSASELEAEHFTFP 240
Db 181 RTFGLPTVQLAYLRTESWRNQAYVATEATTTSAEATTPPTVATTSASELEAEHFTFP 240
QY 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGVTVAAVVSATIGLVIVISIVTNNMCTP 300
Db 241 WLENGVDHYEPTPANENSNVTVRLGTMSPTLIGVTVAAVVSATIGLVIVISIVTNNMCTP 300
QY 301 HRKLDVTSQDDERSOTRESRKFGPMVACEINKGADQDSELVELVAIVNPSALSPDSI 360
Db 301 HRKLDVTSQDDERSOTRESRKFGPMVACEINKGADQDSELVELVAIVNPSALSPDSI 360
QY 361 KM 362
Db 361 KM 362

RESULT 2
O39309 PRELIMINARY; PRT; 420 AA.
AC O39309;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Counterpart of HSV-1 gene US7 and VZV gene 67.
GN 73.
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NS80567;
RX MEDLINE=98264497; PubMed=9603335;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-4.";
RL J. Gen. Virol. 79:1197-1203 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NS80567;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030027; AAC59593.1; -.
DR PIR; T42616; T42616.
DR InterPro; IPR002874; Herpes_g1.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF01688; Herpes_g1; 1.
SQ SEQUENCE 420 AA; 45710 MW; 1EB944825DF3D852 CRC64;

Query Match 8.8%; Score 164; DB 12; Length 420;
Best Local Similarity 25.5%; Pred. No. 2.1e-06;
Matches 96; Conservative 50; Mismatches 143; Indels 88; Gaps 21;

QY 10 LLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREAIQI-QLFPMGQR-PHKPYS 67
Db 13 LLAISMCM--ATAIYRGEHMSMYLNASSSEFAVYPKDKSLVYVGHMLFDGRLPTNYS 69
QY 68 GTVAVAFRSDITNQCQELSEERFNCNTHRSSSVFVGC--KYTEYTFASNRLTGPHPF 125
Db 70 GLIEL-IHNYSRGCVIQTISYSCFRVANNAFRSLHKNHNDQVFNVTNVTN 128
QY 126 KLTIRNPRNDGMYFVIVRLDDTKEPIDVFAIQLSVYQF-ANTAATRGLYSKASCTFG 184
Db 129 LLNITRPOPADSGAIIIRKLNHA-PTADVFGSAFVLDIQSNVPEPPTAKEPSNVT 187
QY 185 LPTVOLAYLRTESWRNQAYVATEATTTSAEAT--TPTPVATTSASELEAEHFTFP 241
Db 188 RTPAPANTSK-----TGSNTTSSQSTWLYTTP-----RPA 221
```

```
QY 242 LENGVDHYEPTPANEN--SNVTVRL-----GTMSET-----LIGTVAAVVSATIG----- 285
Db 222 LET--HLTTPANETVWSGDTAMLCHGRPSTAVPTIYMHLLGLT-----GNLPE 269
QY 286 --LVIVISIVTNNMCTPHKLDVTS-----ODDEERSQTERESF-KFGPMV-----ACEIN 333
Db 270 DVLLIEDSEILR---TPPPKQPTTSRTGDDPKQTNSTSPKSRNKIVAMVVIPTACVLM 326
QY 334 KGADQDSELVELVAIVN 350
Db 327 -----LLLVVVGAIIN 337

RESULT 3
Q98VN1 PRELIMINARY; PRT; 354 AA.
ID Q98VN1
AC Q98VN1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORE67.
OS Human herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10335;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VZV-32;
RX MEDLINE=21109116; PubMed=11162813;
RA Faga B., Maury M., Bruckner D.A., Grose C.;
RT "Identification and Mapping of Single Nucleotide Polymorphisms in the
RT Varicella-Zoster Virus Genome.";
RL Virology 280:11-6 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VZV-32;
RA Cole N.L., Faga B.P., Grose C.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=VZV-VIA;
RA Grose C., Faga B.;
RT "Identification and mapping of single nucleotide polymorphisms in the
RT varicella-zoster virus genome.";
RL Submitted (DSC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF314221; AAK19249.1; -.
DR EMBL; AF325441; AAK01055.1; -.
DR InterPro; IPR002874; Herpes_g1.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF01688; Herpes_g1; 1.
SQ SEQUENCE 354 AA; 39373 MW; 3F01739F3AFC6B08 CRC64;

Query Match 8.7%; Score 162.5; DB 12; Length 354;
Best Local Similarity 22.4%; Pred. No. 2.2e-06;
Matches 72; Conservative 53; Mismatches 111; Indels 85; Gaps 15;

QY 23 IVITGNHVSARIDDDHIVIVAPRE---ATIQQLFFMPGQR-PHKPYSGTVRVAFRSDI 78
Db 21 LIKFGDHSVLSQVNSLSLILPMQNDNVTIKGLVIGEQLPTGNTVSGTLELLY-ADT 79
QY 79 TNQCQELSEERFNCNTHRSSSVFVGCQVT---BYTFSASNRLTGPHPFKLTIRNPRN 135
Db 80 VAFCFRSVQVIRYDGCPRITSAFISCRYKHSWYGNSTDRISTEPDAGVMKIKTPGIN 139
QY 136 DSGMFPYIVRLDDTKEPIDVFAIQLSVY-----QFANTAATRGLYSKAS 179
Db 140 DAGVYVLLVLDHRS-TDGFILGVNVYTAGSHNHGIVYTSPLQNGYSTRALFQA- 197
QY 180 CRTGELPT-----VQLEAYLRTES-WRNQAYVATEATTTSAEATTPPT 225
Db 198 -RLCDLPATPKSGTSLFQHMULDRLAGKSLDNLFWLH-EDVVTETKSVKEG----- 248
QY 226 ATSASELEAEHFTFPWLENGVDHYEPTPAN-----ENS-----NVTVRLGTSPTLIGTV 276
```

```
Db 249 -----IEN-----HVYPTDMSTLPEKSLNDPENLLI-----LIPIVA 281
QY 277 AAVVSATGLVIVISVTRNM 297
Db 282 SVMILTAMVIVISVTRRI 302

RESULT 4
OS5525 ID O55525 PRELIMINARY; PRT; 420 AA.
AC O55525;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Glycoprotein I homologue.
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OC NCBI_TaxID=10331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TH20;
RX MEDLINE=96185635; PubMed=9524947;
RA Damiani A.M., Matsumura T., Yokoyama N., Maeda K., Miyazawa T.,
RA Kai C., Mikami T.;
RT "Nucleotide sequences of glycoprotein I and E genes of equine
RT herpesvirus type 4";
RL J. Vet. Med. Sci. 60:219-225 (1998).
DR EMBL; AB005749; BAA25021.1; -
DR InterPro; IPR002874; Herpes GI.
DR Pfan; PF01688; Herpes GI.
DR Pfan; PF01688; Herpes GI.
SQ SEQUENCE 420 AA, 45696 MW, 1F64F5E179C2D916 CRC64;

Query Match 8.5%; Score 158; DB 12; Length 420;
Best Local Similarity 25.5%; Pred. No. 7.3e-06;
Matches 96; Conservative 49; Mismatches 144; Indels 88; Gaps 21;

QY 10 LLAATLAPFAGMIVITGNHVSARIDDDHIVIVAPRPEATLQLQLFPMQGR-PHKPYS 67
Db 13 LLAISMC---ATAIYRGHSMYLNASSEFAVYPKDKSLVVVGHMLFDGRLPTNYS 69
QY 68 GTVAVFRSDITNQCQLSEERFENCTHRSSSVFVGC--KVTEYTFASNRLGPPHPF 125
Db 70 GLIEL-IHNYSGYCVISQITTSYSCPRVANNAPRSCRLHKTSHNQDYFHTVSTVETNV 128
QY 126 KLITRNPRNDGMEFYVIVRLDDTKEPIDVFAIQLSVYQF-ANTAAATRGLYSKASCRFTG 184
Db 129 LLNITWFPQADSGAILRVKLNHA-PTADVFGVSFAVYDLSQNTVPEPPTAKEPSNVFT 187
QY 185 LPTVQLEAYLRTEESWRNMQAVATEATTSABAT---TPTPTVATSAASELEAEHFTFPW 241
Db 188 RTPAPAPANTSTK-----TGSNTTSQSSTWLYTTP-----RPA 221
QY 242 LENGVDHYEPTANEN--SNVTVL-----GTMSPT-----LIQVTVAAVVSATIG---- 285
Db 222 LET---HLTAPANETVSGDTAMLCHGRPRPTAVPTIYMLLGLT-----GNLPE 269
QY 286 --LVIVISVTRNMCTPRKLTDSQDDE-----ERSOTRRRES-KFGPMV-----ACIN 333
Db 270 DVLIIEDSEILR---TPPKPQOTTSSRTGDAFKQTNSTSPSKRNKIVAMVVIPTACVIM 326
QY 334 KGADQDSSELVELVAIVN 350
Db 327 -----LLLVVVGAIIIN 337

RESULT 5
Q69357 ID Q69357 PRELIMINARY; PRT; 384 AA.
AC Q69357;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
```

```
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Homologue of HSV-1 GI.
OS Feline herpesvirus (Feline herpesvirus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OC NCBI_TaxID=10334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G2620;
RX MEDLINE=95266277; PubMed=7747442;
RA Willemse M.J., Strijdeven I.G., van Schooneveld S.H.,
RA van den Berg M.C., Sondermeijer P.J.;
RT "Transcriptional analysis of the short segment of the feline
RT herpesvirus type 1 genome and insertional mutagenesis of a unique
RT reading frame.";
RL Virology 208:704-711 (1995).
DR EMBL; D42113; BAA07692.1; -
DR InterPro; IPR002874; Herpes GI.
DR Pfam; PF01688; Herpes GI.
SQ SEQUENCE 384 AA, 43009 MW, 670C0893CEAAB9DC CRC64;

Query Match 8.4%; Score 156; DB 12; Length 384;
Best Local Similarity 23.3%; Pred. No. 9.9e-06;
Matches 75; Conservative 48; Mismatches 141; Indels 58; Gaps 12;

QY 5 LGTLALATLAPFAGMIVITGNHVSARIDDDHIVIVAPRPE-ATLQLQLFPMQGR-RP 62
Db 1 MSSIAFIYILMAIGTVYGVIRGDHVSHTVSSGFIYPTLENFTYHGLIFLDQPLP 60
QY 63 HKPYSGTVRVAFRSDITNQCQLSEERFENCTHRSSSVFVGC--KVTEYTFASNRLTG 120
Db 61 VNNYNGTLEI-IHYNHSSCYKIVQIEYSSCPVRNNAFRSCRLHKTSMHQYDQLSINTS 119
QY 121 PPHPPKLTITNPRNDGMEFYVIVRLDDTKEPIDVFAIQLSVYQFANTAAATRGLYSKASC 180
Db 120 VETGMLTITSPQVEDGGIYALVRFRFNHNNK-ADVFGLSVFVYSP----DTRGRRHAD- 173
QY 181 RTFGLPTVQLEAYLRTEESWRNMQAVATEATTSABATTPPTVATSAASELEAEHFTFP 240
Db 174 -----ENLNGSILLTTPS--SMETVYKVTPIYDHMTTQT-----TSNKSMESE----P 216
QY 241 WLENGVDHYEPTANENSNVTVL-----GTMSPTLIGVTVAAVVSAT- 283
Db 217 SNTSISCHFPQNDPNEGELYTHLLIAGNITVDYDDVMDGTTLQPRLLDGLNLSVTSF 276
QY 284 -----IGLVIVISI 292
Db 277 KNETTQKTWTPDKVGFVIVISI 298

RESULT 6
Q66931 ID Q66931 PRELIMINARY; PRT; 384 AA.
AC Q66931;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Feline herpesvirus type 1 GI gene precursor.
GN GI.
OS Feline herpesvirus (Feline herpesvirus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OC NCBI_TaxID=10334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B927;
RX MEDLINE=96357074; PubMed=8764058;
RA Mijnes J.D.F., van der Horst L.M., van Anken E., Horzinek M.C.,
RA Rottier P.J.M., De Groot R.J.;
RT "Biosynthesis of glycoproteins E and I of Feline herpesvirus: gE-gI
RT interaction is required for intracellular transport.";
RL J. Virol. 70:5466-5475 (1996).
```

DR EMBL; X98448; CAA67077.1; -.
DR InterPro; IPR002874; Herpes_gi.
DR Pfam; PF01688; Herpes_gi; 1.
KW SIGNAL.
FT SIGNAL 1 18 POTENTIAL.
SQ SEQUENCE 384 AA; 43019 MW; C9ACE1429B496DC1 CRC64;

Query Match 8.3%; Score 154; DB 12; Length 384;
Best Local Similarity 22.3%; Pred. No. 1.5e-05;
Matches 73; Conservative 52; Mismatches 133; Indels 70; Gaps 13;

QY 5 LGTLALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRPE-ATIQQLFFMPGQ-RP 62
DB 1 MSIAPIVILMAIGTVYGRGDVSLHVDTSFGVIVPTLENTIYCHLIFLDQPLP 60

QY 63 HKPYSGTVAVRSDITNOCYQELSEERFENCTHRSSSVFVGC--KVTEYFASNLRTG 120
DB 61 VANNYNGTLEI-THYHSHSCYKIVQVIEYSSCPVRNNAFRSCLHKTSMHQYDQLSINTS 119

QY 121 PHPPFKLTIRNPRDPSGMYVIVRLDDTKEPIDVFAIQLSVYQANTATRGLYSKASC 180
DB 120 VETGMLLTITSPKMDGGIYALRVFNHNK-ADVGLSVFYISF-----DTRGHRHDAE 174

QY 181 RTFG--LPTVQ--LEALRTEESRWQAY---VATEATTSAEATTPVPTATSASELEA 234
DB 175 NLNGEILTTPSPMETTVKV-----NTPYDHVNTTQTSKNSESPNTSISC----- 223

QY 235 EHFTFPWLENGVDHYEFTPANENSNVTVRL-----GTMSPTLIGTVVAA 278
DB 224 -----HTQNDPNEGETLYTHLLNAGNTYDDMMVMDGTTLKPLIDWGLNL 270

QY 279 VVSAT-----IGLVIVISI 292
DB 271 SVTSSPKNETQKTPDRKVGIVIVISI 298

RESULT 7
Q9E6L5 PRELIMINARY; PRT; 355 AA.
AC Q9E6L5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glycoprotein I.
GN MDV095 OR GIS.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Md5;
RX MEDLINE=20392152; PubMed=10933706;
RA Tulman E.R., Afonso C.L., Lu Z., Zeak L., Rock D.L., Kutish G.F.;
RT "The genome of a very virulent Marek's disease virus."
RL J. Virol. 74:7980-7988 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Md5;
RA Tulman E.R., Afonso C.L., Lu Z., Zeak L., Rock D.L., Kutish G.F.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=X;
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vv+MDV): Mutations
in the Glycoprotein L-encoding Gene in Some vv+MDVs";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RX EMBL; AF243438; AAG14269.1; -
DR EMBL; AY129990; AAM97723.1; -
DR InterPro; IPR002874; Herpes_gi.
DR Pfam; PF01688; Herpes_gi; 1.

SQ SEQUENCE 355 AA; 40568 MW; D23A6DA126ACDB00 CRC64;

Query Match 8.2%; Score 152; DB 12; Length 355;
Best Local Similarity 22.9%; Pred. No. 2.1e-05;
Matches 78; Conservative 55; Mismatches 166; Indels 42; Gaps 14;

QY 19 GAMGIVITGNHVSARIDDDHIVIVAPRPE-ATIQQLFFMPGQRPKPKPGSGTVVAFRSD 77
DB 15 GIWSIVVTGTSVTLSDQSALVAFCGLDKMNVRGQLFLGDQTRTSSYTGTTET-LKWD 73

QY 78 ITNOCYQELSEERFENCTHRSSSVFVGCVKTEYFASNLRTGPPHPFK---LTIRNRP 134
DB 74 EYKCYSLVATSYMDCPAIDATVFRGCRDAVVYAOPHDRV--OPPEKGTLLRIVEPRV 131

QY 135 NDSGMFYIVRLDDTKEPIDVFAIQLSVYQANTATRGLYSKASCRTFGLPTVQLEAYL 194
DB 132 SDTGSYIIRVALAG-RNMSDIFRMAVIIRS-----SKSWACNHSASSFOAHKCI 179

QY 195 RTEE--SWRNQAYVATEATTSAEATTPVPTATSASELEAHEFTFPWLENGVDHYEPT 252
DB 180 RYVDRMAFENVLGHVGNLLDSDSELHAYINIPQGIS-TDINIITPFYDNGSTIYSPT 238

QY 253 P---ANENSNVTVRLGT-MSPTLIGTVVAAV--SATIGLVIV---ISIVTRNMCTPHRK 303
DB 239 VFNLFNNSHVDAMNSTGMWNTVLYKTLPLRIYFSTMIVLICIIALAIYLCERCSPHRR 298

QY 304 LDTVSQDDERSQTRRESKFGPMVACEINKGADQDSOLVE 344
DB 299 I---YIGERSDE-----APLIITSANVESFOYDYNWKE 328

RESULT 8
Q8JUL8 PRELIMINARY; PRT; 355 AA.
AC Q8JUL8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein I.
GN GIS.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=N;
RA Parcells M.S., Shamblin C.E., Dienglewicz R.L.;
RT "DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vv+MDV): Mutations
in the Glycoprotein L-encoding Gene in Some vv+MDVs";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RX EMBL; AY129992; AAM97725.1; -
DR InterPro; IPR002874; Herpes_gi.
DR Pfam; PF01688; Herpes_gi; 1.
SQ SEQUENCE 355 AA; 40584 MW; 58DB07B04193C0A0 CRC64;

Query Match 8.1%; Score 151; DB 12; Length 355;
Best Local Similarity 23.2%; Pred. No. 2.6e-05;
Matches 79; Conservative 53; Mismatches 167; Indels 42; Gaps 14;

QY 19 GAMGIVITGNHVSARIDDDHIVIVAPRPEATIQQLFFMPGQRPKPKPGSGTVVAFRSD 77
DB 15 GIWSIVVTGTSVTLSDQSALVAFCGLDKMNVRGQLFLGDQTRTSSYTGTTET-LKWD 73

QY 78 ITNOCYQELSEERFENCTHRSSSVFVGCVKTEYFASNLRTGPPHPFK---LTIRNRP 134
DB 74 EYKCYSLVATSYMDCPAIDATVFRGCRDAVVYAOPHDRV--OPPEKGTLLRIVEPRV 131

QY 135 NDSGMFYIVRLDDTKEPIDVFAIQLSVYQANTATRGLYSKASCRTFGLPTVQLEAYL 194
DB 132 SDTGSYIIRVALAG-RNMSDIFRMAVIIRS-----SKSWACNHSASSFOAHKCI 179

Qy	195	RTEE--SWRNQAYVAT	EATTTSAAETPTVTATSASELEAHFTFPWLENGVDHYEPT	255
Dd	180	RYVDRMAFENYLIGHVGNLSDSELHAIYNIITPQSIS-TDINIITTPFDNSGTIYSPT	238	
Qy	253	P---ANENSNTVRLGT-KSTPLIGTVAAVV--SATIGLVIV---	ISIIVTRNKCTPHRK	303
Dd	239	VFNLFNNSHVDAMNMGWNVTKYTLPRVIYESTMVICIIALAIYLCERCSPPHR	298	
Qy	304	LDTVSQDEESQTRRESKFGPMVACINKGADQDSOLVE	344	
Dd	299	I-----YIGEPSDE-----APLTSAVNSESFOYDNVKE	328	
 RESULT 9 Q69287				
ID	Q69287	PRELIMINARY;	PRT; 355 AA.	
AC	Q69287;			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)		
DT	01-OCT-2002	(TREMBLrel. 22, Last annotation update)		
DE	Glycoprotein I.			
GS	US7 OR GIS.			
ON	Turkey herpesvirus.			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_TaxId=10390;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95133166; PubMed=7831788;			
RA	Brunovskis P. Velicer L.F.;			
RT	"The Marek's disease virus (MDV) unique short region:			
RT	alphaherpesvirus-homologous, fowlpox virus-homologous, and MDV-			
RT	specific genes.";			
RL	Virology 206:324-338(1995).			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=RBLB;			
RA	Parcells M.S., Shamblin C.E., Dienglewicz R.L.;			
RT	"DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three			
RT	Pathotypes of Marek's Disease Viruses (vMDV, vvMDV, vvvMDV): Mutations			
RT	in the Glycoprotein L-encoding Gene in Some vv+MDVs.";			
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
DL	EMBL; L22174; AAA64968.1; -.			
DR	EMBL; AV129994; AAM97727.1; -.			
DR	InterPro; IPR002874; Herpes_GI.			
DR	Fram; PF01688; Herpes_GI; 1.			
KX	Signal.			
FT	SIGNAL	1 18 POTENTIAL.		
FT	CHAIN	19 355 MEMBRANE GLYCOPROTEIN I.		
SQ	SEQUENCE	355 AA; 40579 MW; 40CD25034E4EA6GF CRC64;		
 Query Match 8.0%; Score 149; DB 12; Length 355; Best Local Similarity 23.3%; Pred. No. 3.9e-05; Matches 80; Conservative 54; Mismatches 163; Indels 46; Caps 15				
Qy	19	GANGVITGNHVSARIDDDHVIVAVR---	PEATIQLFLPMGPQRHPKPYSGTVAVR	75
Dd	15	GIWSIVYTGSTVT--LSTDQSAIAFAFRGLDKVANVRGQLFLGDQTFTSSYTGTTEI-LK	71	
Qy	76	SDITWCYQELSEEPENCETHRSSVFVGCKVTEVTSASNLRTGGPPHPFK--	LLIRNP	132
Dd	72	WBEEYKCYSVLHATSYMDCPAIDATVFRGCRDAVVVYAQPGRGV--	QPFPEKGLILLRIIVEP	129
Qy	133	RPNDSCGMFYIVRLDDTEKDIPVAIQLSVQFANTAAATRGLYSKASCRTFGLTVPQLEA	192	
Dd	130	RVSDTGSYIRVSLAG-RNMSDIFRMVIIRS-----	SKSWACHNSASSFOAKH	177
Qy	193	YLRTEE--SWRNQAVVATEATTTSAAETPTVTATSASELEAHFTFPWLENGVDHYE	250	
Dd	178	CIRYVDRMAFENYLIGHVGNLSDSELHAIYNIITPQSIS-TDINIITTPFDNSGTIYS	236	
Qy	251	PTP---ANENSNTVRLGT-MSPTLIGTVAAVV--SATIGLVIV---	ISIIVTRNKCTPH	301

```
RC STRAIN=YPI1LMU;
RX MEDLINE=98455388; PubMed=9784067;
RA Nishikawa Y., Xuan X., Otsuka H.;
RT "Identification and characterization of the glycoprotein E and I genes
RT of canine herpesvirus.";
RN Virus Res. 56:77-92(1998).
RP [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Australian;
RA Reubel G.H., Pekin J., Webb-Wagg K., Hardy C.M.;
RT "Nucleotide sequence of glycoprotein genes B, D, G, H and I, thymidine
RT kinase and protein kinase genes and gene homologue UL 24 of an
RT Australian isolate of canine herpesvirus.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U84223; AAB67059.1;
DR EMBL; AF003729; BAA33765.1;
DR EMBL; AF361076; AAK51063.1;
DR GO; GO:0005529; F: sugar binding; IEA.
DR InterPro; IPR002874; Herpes_gi.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01688; Herpes_gi; 1.
DR SMART; SMO0034; CLECT; 1.
SQ SEQUENCE 364 AA; 42068 MW; 08BC32EF7FB913A8 CRC64;

Query Match 7.9%; Score 148; DB 12; Length 364;
Best Local Similarity 27.3%; Pred. No. 5e-05;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQQLFFMPQGR-P 62
Db 16 LLTMTFLPILFLYGVNGFYKGTIYSMLFNSSGFSIPDDKFIYSGRLFLDDQHL 75
QY 63 HKPYSGTVRAFRSDITNQCYQLSEERFENCNTHRSSSVFVGC--KVTEYTFASNRLTG 120
Db 76 VNNYSGTIEFI---HFNNSCYTVYQTEYFSCPRIFNNAFRSCLKVKSHESQLRINS 132
QY 121 PPHPKLTIRNPRNDGMPFYIVRLDDTKPEIDVFAIQLSVYQF 165
Db 133 IENGVLLEITNPKNDGSGVFIRVQLENNK--TDVFGIPAFIYSF 175
SQ SEQUENCE 370 AA; 41568 MW; 2A816ECB37466A77 CRC64;

Query Match 7.8%; Score 145; DB 12; Length 370;
Best Local Similarity 22.0%; Pred. No. 9.6e-05;
Matches 72; Conservative 52; Mismatches 136; Indels 68; Gaps 14;

QY 5 LGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRE-ATTQQLFFMPQGR-P 62
Db 1 MSSIAFIYILMAIGTVYGVYRGDVSLSHVDTSSTGFIYPTLENFTIYGHILFDQDPLP 60
QY 63 HKPYSGTVRAFRSDITNQCYQLSEERFENCNTHRSSSVFVGC--KVTEYTFASNRLTG 120
Db 61 VNNYNGTLEI-IHYNHSSCYKIVQVIEYSSCPVRNNAFRSCLHKTSMHOYDQLSINTS 119
QY 121 PPHPKLTIRNPRNDGMPFYIVRLDDTKPEIDVFAIQLSVYQFANTAARGLYSKASC 180
Db 120 VETGMLLTITSPKMDGGIYALRVFRFNHNK-ADVFGLSVFVYSF---DTRGHRHHADE 174
QY 181 RTFG--LFTVQ--LEAYLFTESWRNQQAY---VATEATTTSAEATTPFPVATASALEA 234
Db 175 NLANGILLTPSPMETYKVV-----NTPYIDEMVTTQTSNKSMESEPTNTSISC----- 223
QY 235 EHTFPWLENGVDHYEPTPANENSVTVRL-----GTMSFTLIGTVAA 278
Db 224 -----HTFQNDPNEGETLYTHLLINAGNITYDDMVMMDGTTLKPRLLI----- 264
QY 279 VVSATIGLVIVISIVTRNNCTPHRKLDT 306
Db 265 ----DMGLNLSVTSPFNK--GNHAKQDT 286

RESULT 14
QYULU7
ID Q8JLU7 PRELIMINARY; PRT; 355 AA.
AC Q8JLU7;
```

```
RC STRAIN=YPI1LMU;
RX MEDLINE=98455388; PubMed=9784067;
RA Nishikawa Y., Xuan X., Otsuka H.;
RT "Identification and characterization of the glycoprotein E and I genes
RT of canine herpesvirus.";
RN Virus Res. 56:77-92(1998).
RP [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Australian;
RA Reubel G.H., Pekin J., Webb-Wagg K., Hardy C.M.;
RT "Nucleotide sequence of glycoprotein genes B, D, G, H and I, thymidine
RT kinase and protein kinase genes and gene homologue UL 24 of an
RT Australian isolate of canine herpesvirus.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U84223; AAB67059.1;
DR EMBL; AF003729; BAA33765.1;
DR EMBL; AF361076; AAK51063.1;
DR GO; GO:0005529; F: sugar binding; IEA.
DR InterPro; IPR002874; Herpes_gi.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01688; Herpes_gi; 1.
DR SMART; SMO0034; CLECT; 1.
SQ SEQUENCE 364 AA; 42068 MW; 08BC32EF7FB913A8 CRC64;

Query Match 7.9%; Score 148; DB 12; Length 364;
Best Local Similarity 27.3%; Pred. No. 5e-05;
Matches 45; Conservative 30; Mismatches 82; Indels 8; Gaps 4;

QY 4 LLGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPREATIQQLFFMPQGR-P 62
Db 16 LLTMTFLPILFLYGVNGFYKGTIYSMLFNSSGFSIPDDKFIYSGRLFLDDQHL 75
QY 63 HKPYSGTVRAFRSDITNQCYQLSEERFENCNTHRSSSVFVGC--KVTEYTFASNRLTG 120
Db 76 VNNYSGTIEFI---HFNNSCYTVYQTEYFSCPRIFNNAFRSCLKVKSHESQLRINS 132
QY 121 PPHPKLTIRNPRNDGMPFYIVRLDDTKPEIDVFAIQLSVYQF 165
Db 133 IENGVLLEITNPKNDGSGVFIRVQLENNK--TDVFGIPAFIYSF 175
SQ SEQUENCE 370 AA; 41568 MW; 2A816ECB37466A77 CRC64;

Query Match 7.8%; Score 145; DB 12; Length 370;
Best Local Similarity 22.0%; Pred. No. 9.6e-05;
Matches 72; Conservative 52; Mismatches 136; Indels 68; Gaps 14;

QY 5 LGTALLAATLAPFGAMGIVITGNHVSARIDDDHIVIVAPRE-ATTQQLFFMPQGR-P 62
Db 1 MSSIAFIYILMAIGTVYGVYRGDVSLSHVDTSSTGFIYPTLENFTIYGHILFDQDPLP 60
QY 63 HKPYSGTVRAFRSDITNQCYQLSEERFENCNTHRSSSVFVGC--KVTEYTFASNRLTG 120
Db 61 VNNYNGTLEI-IHYNHSSCYKIVQVIEYSSCPVRNNAFRSCLHKTSMHOYDQLSINTS 119
QY 121 PPHPKLTIRNPRNDGMPFYIVRLDDTKPEIDVFAIQLSVYQFANTAARGLYSKASC 180
Db 120 VETGMLLTITSPKMDGGIYALRVFRFNHNK-ADVFGLSVFVYSF---DTRGHRHHADE 174
QY 181 RTFG--LFTVQ--LEAYLFTESWRNQQAY---VATEATTTSAEATTPFPVATASALEA 234
Db 175 NLANGILLTPSPMETYKVV-----NTPYIDEMVTTQTSNKSMESEPTNTSISC----- 223
QY 235 EHTFPWLENGVDHYEPTPANENSVTVRL-----GTMSFTLIGTVAA 278
Db 224 -----HTFQNDPNEGETLYTHLLINAGNITYDDMVMMDGTTLKPRLLI----- 264
QY 279 VVSATIGLVIVISIVTRNNCTPHRKLDT 306
Db 265 ----DMGLNLSVTSPFNK--GNHAKQDT 286

RESULT 14
QYULU7
ID Q8JLU7 PRELIMINARY; PRT; 355 AA.
AC Q8JLU7;
```

DT	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
GN	Glycoprotein I.
OS	Turkey herpesvirus.
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	Alphaherpesvirinae; Marek's disease-like viruses.
NCBI_TaxID=10390;	
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=PK;
RA	Parcells M.S., Shamblin C.E., Dinglewicz R.L.;
RT	"DNA Sequence Analysis of Glycoprotein-Encoding Genes of Three
RT	Pathotypes of Marek's Disease Viruses [vMDV, vVMVD, vv-MDV]; Mutations
RT	in the Glycoprotein L-encoding Gene in Some vv-MDVs.";
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AV129993; AM97726.1; -
DR	InterPro; IPR002874; Herpes gl. 1.
DR	Pfam; PF01688; Herpes gl. 1.
DR	PFam; PF01688; Herpes gl. 1.
SQ	SEQUENCE 355 AA; 40599 MW; 316993FF97143C1E0 CRC64;

Query Match	7.7%; Score 144; DB 12; Length 355;
Best Local Similarity	22.6%; Pred. No. 0.00011;
Matches 77; Conservative 55; Mismatches 167; Indels 42; Gaps 14;	

Qy	19	GAMGIVITGNHVSARIDDDHHVIIVAPRPE-ATQLQLFMPGORPHKPYSGTVRVAFPSD	77
		:	:
Db	15	GIWSRVYTGTSVTLTQSALVAFGLDKMVRGQLFLFGDQTRTSSYTGTET-LKWD	73
		:	:
Qy	78	INNOCYQELSEERFENCSTRHSVVGVCKKYTEFTFSASNLTCPPHPFK---LTIRNP	134
		:	:
Db	74	EYKCVSLIHATSVMDCPAIDAIVFRGCRDAVVYAOPHDRV--QPFPKGTLLRIIVEPRV	131
		:	:
Qy	135	NDSGMFYIVLRDDTKEPIDYFAQLSVYQFANTATRGLYSKASCRTFGLPTVOLAEVL	194
		:	:
Db	132	SDTGSYIRVALAG-RNMSDFRWAVIIRS-----SKSWACNSASSFOAHKCI	179
		:	:
Qy	195	RYTEE--SWRNMQAYVATEATTTSABATPPTVATSAASELEASHFTFPWLENGVDHYBPT	252
		:	:
Db	180	RYVDMRAFENYLIGHVNLLDSDELHAIYNIPQIS-TDINIITPPFYDNSGTIYSPT	238
		:	:
Qy	253	P---ANENSNTVRLGT-MSPTLGIVTAAVV--SATIGLVIV---ISIVRMCTPHRK	303
		:	:
Db	239	VFNLFNNSHVDAMNSTGMATVLTLYTLPLIIYSTMTVICITAIALIVLCERCSPHR	298
		:	:
Qy	304	LDTVSDDESRQTRRESKFGPMVACEINKGADQDSOLVE	344
		:	:
Db	299	----TVIGEPRDE-----APLITSAVNESFOYDYNVKE	328
		:	:

RESULT 15	
Q91E39	PRELIMINARY; PRT; 369 AA.
ID	Q91E39
AC	Q91E39;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
GN	Glycoprotein I.
DE	US?
OS	Phocid herpesvirus 1.
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	Alphaherpesvirinae; Varicellovirus.
NCBI_TaxID=47418;	
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=PB84;
RA	Martina B., Osterhaus A.D.M.E., Harder T.C.;
RT	"Identification and analysis of immunogenicity of the glycoprotein D
RT	equivalent within the unique short segment of phocid herpesvirus-1.";
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AJ290955; CAC51466.1; -
DR	InterPro; IPR002874; Herpes gl.

```
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF01688; Herpes_gI; 1.  
SQ SEQUENCE 369 AA; 42372 MW; 1587181704A7AA8A CRC64;  
  
Query Match  
Best Local Similarity 25.1%; Score 140.5; DB 12; Length 369;  
Matches 57; Conservative 36; Mismatches 89; Indels 45; Gaps 10;  
  
QY 22 GIIVTGNHVSARIDDDHIVIVAPRP--EATIQQLFFMPGQR-PHKPYSGTVRVAFRSBIT 79  
||| : : : : :  
Dd 16 GIIVRGTYMSMYNTSSGIYVDPDDRDFNVGYLLFLDQRLPVNYSGTIIYY---FN 72  
||| : : : : :  
  
QY 80 NCQCYQELSEERFENCSTRSSSVFGC--KVTEYTFSA-----SNRLTGPPHPFKLTINR 133  
||| : : : : :  
Dd 73 YSCYTVVTQYTEYVSCPIRHNAFRSCLIKVKSKHQSLRINSIETG---VLLEIKNPK 128  
||| : : : : :  
  
QY 134 PDSGMFYIVRLDDTKBPDVFALISVYQFANTAATGLYSKASCRTEGLPTVQLEAY 193  
||| : : : : :  
Dd 129 PSDSGVIYPRVQLENKK--TDVGISAFVYSPFKSG--ENITKPDNSOT----- 173  
||| : : : : :  
  
QY 194 LRTEBSRWNRQAQVAATEATTTSABA-----TTPTPVTATASEL 232  
||| : : : : :  
Dd 174 -----ENFTLVLPSTISTSKPFSETSHLMTFPFDIPAPYCHEV 213  
||| : : : : :
```

Search completed: March 8, 2004, 02:44:33
Job time : 128 secs

```

RESULT 15
Q91E39
ID Q91E39 PRELIMINARY; PRT; 369 AA.
AC Q91E39;
DT 01-DEC-2001 (trEMBLrel. 19, Created)
DT 01-DEC-2001 (trEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (trEMBLrel. 25, Last annotation update)
DE Glycoprotein 1.
DE US7.
OS Phocid herpesvirus 1.
OS Viruses; dsDNA viruses, not RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=47418;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PB84;
RA Martina B., Osterhaus A.D.M.E., Harder T.C.;
RT Identification and analysis of immunogenicity of the glycoprotein D
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ290955; CAC51466.1; -
DR InterPro; IPR002874; Herpes g1.
DR

```